Ambiguous Perception and Selective Attention - Competitive Processes in Complex Scenarios

Dissertation

zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Physik
der Philipps-Universität Marburg
vorgelegt von

Svenja Marx
aus Giessen

Marburg/Lahn 2015

Vom Fachbereich Physik der Philipps-Universität als Dissertation angenommen

am: 25.03.2015

Erstgutachter: Prof. Dr. Wolfgang Einhäuser-Treyer (Universität Marburg)

Zweitgutachter: Prof. Dr. Peter König (Universität Osnabrück)

Tag der mündlichen Prüfung: 17.04.2015

Hochschulkennziffer 1180

Contents

Summary

Our visual system is confronted with complex and often ambiguous scenes and events every day. This information needs to be transduced, filtered and processed to enable us to behave adequately in our environment. This process of interpreting our environment based on information contained in visible light is called visual perception. The challenges our visual perception is confronted with are manifold. For example, lack of information makes situations difficult to interpret and sometimes visual scenes may allow different interpretations making it hard to compose congruent experiences from them. Already Helmholtz realized that our sensory system provides us with fragmentary information and that congruent experiences can only be inferred from these fragments ("unconscious inferences"; Helmholtz, 1867). The way this is usually done is that the most likely interpretation is chosen based on our expectations derived from prior experience.

The processing of visual information that ultimately leads to perception is thus not exclusively driven by processes reflecting the present sensory stimulus (so-called bottom-up processes), but it is also influenced by higher functions like experience or attention (top-down processes). This thesis consists of five studies. Three of them investigate how bottom-up and top-down processes influence perception in situations challenging the visual system and what the underlying circuitry could be. The final two studies investigate how visual perception changes with age and how eye tracking can be employed as a diagnostic tool.

Attention influences perception as well as other perception-related processes such as learning and searching. Value (i.e., reward or punishment) strongly influences selective attention and decision-making, but whether there is a direct effect of value on perception per se is still unknown. We investigated the influence of reward and punishment on perception using a phenomenon called binocular rivalry. This is a visually ambiguous situation where both eyes are presented with dissimilar stimuli. In this study, we used differently oriented and colored gratings drifting in different directions. Since the visual system cannot merge both stimuli into one percept, only one of the stimuli is perceived at a time and perception switches between the two possible interpretations. By rewarding or experimentally drawing attention to one of the stimuli, we modified different top-down influences on perception. Eye movements were measured throughout the experiments to objectively assess which of the gratings the observer perceived at any given point in time. We found that, when investigated in isolation, attention and reward had similar effects on perception, but, importantly, when investigated simultaneously, there was an additional effect of reward on top of the attentional one. Even though

reward could still be mediated by attention, we could conclude that reward affects perception, even when attention is engaged in a different task.

In a second study, to get a deeper understanding of the processes underlying perceptual alternations observed in binocular rivalry, we developed a computational model for this phenomenon. Although there have been many attempts to model rivalry and all its characteristics in the past, no model has so far succeeded in covering all the hallmarks of rivalry. Motivated by the strong influence of attention on rivalry observed in our first study and because rivalry is sometimes related to processes of visual attention, we chose to use a winner-take-all (WTA) model, which is often used to model attentional processes. Our firing-rate model receives noisy input and consists of two WTA circuits - the input and the perception circuit - each comprised of two excitatory and one inhibitory unit. These circuits were coupled such that, apart from alternations between dominant units, also memory was included. In addition to our modeling approach, we conducted psychophysical experiments and compared the experimental and modeling results directly. With the WTA model we succeeded in replicating all psychophysical findings qualitatively and most of them quantitatively.

Another approach of testing perception apart from ambiguous stimulation is to present stimuli briefly and in rapid succession. This procedure is called rapid serial visual presentation (RSVP) and is an established method to investigate the effects of attention and stimulus features on perception. When two targets are present in an RSVP stream and both appear in close temporal proximity, a phenomenon termed the "attentional blink" can be observed. Perception of the second target is impaired, so it is less frequently detected than without the first one. In our experiment, we were particularly interested in the effect of color on natural-scene processing, since the function of color vision is often associated with attentional processes. We presented natural scenes in an RSVP task including different numbers of targets (animals), which had to be detected and categorized by human observers. To challenge perception to different degrees, we used presentation durations (stimulus onset asynchronies, SOAs) between 30 and 120ms. In addition, we manipulated color of the natural scenes in four conditions: original color, modified color, gray-scale and inverted gray-scale. Color seemed to have different roles for different presentation durations: For short (<=60ms) SOAs color had no influence on performance, for intermediate SOAs (90ms-100ms) there was a benefit for original color images and for the longest SOA (120ms), both color conditions were indistinguishable. This led us to the conclusion that the role of color changes with presentation duration from being of no benefit through being characteristic of the target up to serving figure-ground segmentation. We also observed the attentional blink, but differences between color conditions were fully explained by single-target differences. Furthermore, categorization-given-detection performance was not influenced by color. These two findings led us to the conclusion that color benefits cannot be purely attentional but must also be perceptual.

The three aforementioned studies were all conducted in a laboratory setting. Even though complex and naturalistic stimuli were used, the well-defined laboratory environment was certainly a simplification of the real world's complexity. So we went one step further and measured eye movements in the real world using a wearable eye tracker. Since eye movements are a good indicator of where attentional resources are allocated in space, eye tracking served as a method to measure visual attention. Additionally, eye-movement parameters are a good measure to compare different participant groups. Two studies were conducted using eye tracking in the real world. In the first study, the differences in visual perception between younger and older adults were investigated in a real-world setting. In the second study, we validated the use of mobile eye tracking as a diagnostic tool to discriminate between two neurodegenerative diseases, namely idiopathic Parkinson's disease (IPD) and progressive supranuclear palsy (PSP). In both studies, participants were walking down a corridor during the experiment such that, in addition to eye movements, also body and head movements were executed. In the first study, older adults showed slower saccade mean and peak velocities, shorter saccade amplitudes, and a tendency to execute fewer saccades as compared to young adults. This supports laboratory findings that older adults have a narrower field of view. In the second study, a significant impairment of saccade velocity and amplitude, most prominent in the vertical plane, was observed for PSP patients as compared to IPD patients and healthy controls, which confirmed laboratory findings. Interestingly, the impairment was more prominent in a standardized fixation task we measured in addition to the real-world walking condition, indicating that patients can compensate for their deficits in real life. Taken together, these two studies establish real-world mobile eye-tracking as a quick and easily applicable research method and diagnostic tool.

In conclusion, this thesis sheds light on how perception is influenced by top-down processes like attention and valuation, how perceptual ambiguity is likely to be processed in the brain, and how visual perception changes with age and with neurodegenerative diseases like PSP. The findings support the hypothesis of perception as an active process, which is not only reflex-like driven by the stimulus but also based on expectations and attention. Beyond basic research, our findings provide a first step towards applied fields by parameterizing healthy aging and providing a tool for clinical diagnosis.

Zusammenfassung

Unser visuelles System wird jeden Tag mit komplexen und mehrdeutigen Szenen und Ereignissen konfrontiert. Diese Informationen müssen weitergeleitet, gefiltert und verarbeitet werden, um uns ein angemessenes Verhalten in unserer Umwelt zu ermöglichen. Unter "visueller Wahrnehmung" verstehen wir diesen Prozess der Interpretation auf der Basis von Informationen, die im sichtbaren Licht enthalten sind. Die Herausforderungen, denen sich unsere Wahrnehmung stellen muss, sind vielfältig. Zum Beispiel erschweren fehlende Informationen die Interpretation von Situationen und das Erlangen einer kohärenten Sinneserfahrung, insbesondere da ein und dieselbe visuelle Szene oftmals verschiedene Interpretationen zulassen kann. Schon Helmholtz erkannte, dass uns unsere sensorischen Systeme nur mit bruchstückhaften Informationen versorgen und dass kongruente Erfahrungen nur von diesen Fragmenten abgeleitet werden können ("unbewusste Schlüsse", Helmholtz 1867). Dies geschieht meist, indem die wahrscheinlichste Interpretation gewählt wird, resultierend aus unseren Erwartungen, die auf der Basis vorheriger Erfahrungen entstanden sind.

Der Verarbeitungsprozess von visuellen Informationen, der letztendlich zur Wahrnehmung führt, ist somit nicht ausschließlich von Prozessen, die auf dem sensorischen Stimulus beruhen (sogenannte bottom-up Prozesse), bestimmt, sondern wird ebenso von höheren Funktionen wie Erfahrung und Aufmerksamkeit beeinflusst (top-down Prozesse). Diese Dissertation besteht aus fünf Studien. Drei davon untersuchen, wie bottom-up und top-down Prozesse die Wahrnehmung in für das visuelle System herausfordernden Situationen beeinflussen und welches die zugrundeliegenden neuronalen Mechanismen sein können. Zwei weitere Studien untersuchen, wie sich visuelle Wahrnehmung mit dem Alter verändert und wie Augenbewegungsmessungen zu Diagnosezwecken eingesetzt werden können.

Aufmerksamkeit beeinflusst Wahrnehmung sowie andere wahrnehmungsbezogene Prozesse wie Lernen und visuelle Suche. Der Wert eines Reizes – ob also Belohnung oder Bestrafung darauf folgt – wiederum beeinflusst selektive Aufmerksamkeit und Entscheidungsfindung. Es ist jedoch noch nicht bekannt, ob es einen direkten Effekt von Wertigkeit auf Wahrnehmung gibt. Wir untersuchten den Einfluss von Belohnung und Bestrafung auf die Wahrnehmung mittels eines Phänomens namens "binokulare Rivalität". Dieses Phänomen tritt auf, wenn beiden Augen verschiedene Reize präsentiert werden. In dieser Studie nutzten wir dafür unterschiedlich orientierte und gefärbte Gitter, die sich in verschiedene Richtungen bewegten. Da das visuelle System diese beiden Reize nicht zu einer kohärenten Empfindung zusammensetzen kann, wird zu jedem gegebenen

Zeitpunkt nur ein Reiz wahrgenommen und die Wahrnehmung wechselt zwischen den beiden möglichen Interpretationen. Mit dem Einsatz von Belohnung bzw. der experimentellen Lenkung der Aufmerksamkeit auf einen Reiz, konnten wir die verschiedenen top-down Einflüsse auf Wahrnehmung modulieren. Augenbewegungen wurden während des gesamten Experiments aufgezeichnet, um zu jedem Zeitpunkt objektiv beurteilen zu können, welches der sich bewegenden Gitter der Beobachter wahrgenommen hat. Wir fanden heraus, dass Belohnung und Aufmerksamkeit die Wahrnehmung gleichermaßen beeinflussten, wenn sie getrennt voneinander untersucht wurden. Wenn wir sie jedoch simultan untersuchten, gab es einen zusätzlichen Effekt von Belohnung bzw. Bestrafung zum Effekt der Aufmerksamkeit. Obwohl Belohnung weiterhin durch Aufmerksamkeitsprozesse vermittelt werden könnte, erlauben unsere Ergebnisse den Schluss, dass Belohnung sogar dann Wahrnehmung beeinflusst, wenn die Aufmerksamkeit an eine andere Aufgabe gebunden ist.

In der zweiten Studie entwickelten wir ein Modell für binokulare Rivalität, um ein tieferes Verständnis für die zugrundeliegenden Mechanismen der Wahrnehmungsveränderungen, die bei diesem Phänomen beobachtet werden können, zu erlangen. Obwohl in der Vergangenheit schon viele Versuche unternommen wurden, binokulare Rivalität mit allen dazugehörigen Charakteristiken zu modellieren, hat noch keines dieser Modelle erfolgreich alle Eigenschaften binokularer Rivalität replizieren können. Motiviert durch den in der ersten Studie gefundenen starken Einfluss von Aufmerksamkeit auf binokulare Rivalität, welche häufig mit Prozessen selektiver Aufmerksamkeit in Verbindung gebracht wird, wählten wir ein Winner-take-all (WTA) Modell. Dieses wird häufig für die Modellierung von Aufmerksamkeitsprozessen verwendet. Unser auf neuronalen Feuerraten basierendes Modell erhält ein verrauschtes Eingangssignal und besteht aus zwei WTA Schaltkreisen – dem Eingangs- und dem Wahrnehmungsschaltkreis – wobei jeder wiederum aus zwei erregenden und einer hemmenden Einheit besteht. Diese Schaltkreise wurden so gekoppelt, dass abgesehen von dem Wechsel zwischen den dominierenden Einheiten auch Gedächtniseffekte auftreten. Neben unserer Modellierung haben wir auch psychophysische Experimente durchgeführt und deren Ergebnisse mit den Resultaten der Modellierung direkt verglichen. Mit dem WTA Modell konnten wir erfolgreich alle psychophysischen Ergebnisse qualitativ und in den meisten Fällen auch quantitativ replizieren.

Eine weitere Herangehensweise zur Untersuchung der Wahrnehmung, neben der Präsentation von mehrdeutigen Reizen, ist die sehr kurze Präsentation von Bildern in schneller Abfolge. Diese Methode wird Rapid Serial Visual Presentation (RSVP) genannt und ist unter anderem eine weitere etablierte Methode, um Aufmerksamkeitseffekte sowie Effekte von Stimuluseigenschaften auf Wahrnehmung zu untersuchen. Wenn zwei Zielreize in einer RSVP Bildsequenz enthalten sind und beide in kurzem zeitlichen Abstand erscheinen, kann ein Phänomen beobachtet werden, das „Aufmerksamkeitsblinzeln“ (AB) genannt wird. Dann ist die Wahrnehmung des zweiten Zielreizes beeinträchtigt, er wird also weniger häufig detek-

tiert als in Bildsequenzen ohne den ersten Zielreiz. In unserem Experiment waren wir besonders an dem Effekt von Farbe auf die Verarbeitung von natürlichen Szenen interessiert, da das Farbsehen häufig mit Aufmerksamkeitsprozessen in Verbindung gebracht wird. Wir präsentierten natürliche Szenen in einem Experiment mit RSVP Bildsequenzen, die eine variierende Anzahl von Zielreizen (Tieren) enthielten. Diese mussten von den Versuchspersonen detektiert und in Unterkategorien eingeordnet werden. Um die Wahrnehmung in unterschiedlichem Maß zu fordern, variierten wir die Präsentationsdauern zwischen 30 und 120ms. Zusätzlich manipulierten wir die Farbeigenschaften der Bilder in vier Bedingungen: Originalfarbe, veränderte Farbe, Graustufen und invertierte Graustufen. Wir fanden heraus, dass Farbe verschiedene Funktionen bei unterschiedlichen Präsentationsdauern innehat: Bei kurzer Präsentationsdauer (<=60ms) hatte Farbe keinen Einfluss auf die Detektionsleistung, bei mittlerer Präsentationsdauer (90-100ms) konnten wir einen Vorteil für Farbbilder feststellen, wenn diese in Originalfarbe präsentiert wurden. Für die längste Präsentationsdauer (120ms) gab es keinen Unterschied zwischen Bildern in Originalfarbe und veränderter Farbe. Daraus schlossen wir, dass sich die Rolle, die Farbe bei der Wahrnehmung spielt, mit der Präsentationsdauer verändert. Während sie bei kurzen Präsentationsdauern keinen Effekt hat, wirkt sie bei mittellangen dadurch, dass sie charakteristisch für den Zielreiz ist, wohingegen sie bei längeren Präsentationsdauern die Segmentation von Figur und Hintergrund unterstützt. Wir beobachteten auch das AB, jedoch wurden Unterschiede zwischen den Farbbedingungen vollkommen durch die Unterschiede in RSVP Bildsequenzen mit nur einem Zielreiz erklärt. Weiterhin war die Leistung in der Kategorisierung bei gegebener Detektion nicht beeinflusst von Farbe. Diese Ergebnisse führten uns zu dem Schluss, dass Vorteile in der Wahrnehmung aufgrund von Farbe nicht ausschließlich aufmerksamkeitsbezogen sind, sondern auch wahrnehmungsbezogen sein müssen.

Die drei zuvor genannten Studien wurden alle unter Laborbedingungen durchgeführt. Obwohl komplexe und natürliche Reize benutzt wurden, war die wohldefinierte Laborumgebung sicherlich eine Vereinfachung gegenüber der Komplexität der realen Welt. Deshalb gingen wir einen Schritt weiter und zeichneten Augenbewegungen in der realen Welt auf, indem wir ein tragbares Augenbewegungsmessgerät benutzten. Da Augenbewegungen ein guter Indikator dafür sind, welchen Orten im Raum Aufmerksamkeitsressourcen zugewiesen werden, diente hier die Augenbewegungsmessung unter anderem als Methode zur Messung visueller Aufmerksamkeit. Zusätzlich sind Augenbewegungen ein gutes Maß zur Unterscheidung von verschiedenen Versuchspersonengruppen. Wir führten zwei Studien durch, bei denen wir Augenbewegungen in der realen Welt aufzeichneten. In der ersten Studie wurden die Unterschiede in visueller Wahrnehmung und von Augenbewegungen in der realen Welt zwischen jungen und älteren Erwachsenen untersucht. In einer weiteren Studie validierten wir den Gebrauch von mobiler Augenbewegungsmessung als diagnostisches Werkzeug in der Unterscheidung von zwei neurodegenerativen Krankheiten, der Progressiven Supranukleären

Blickparese (PSP) und dem idiopathischen Parkinsonsyndrom (IPD). In beiden Studien liefen die Versuchspersonen einen Flur entlang, sodass zusätzlich zu den Augenbewegungen auch Körper- und Kopfbewegungen ausgeführt wurden. In der ersten Studie zeigten ältere Erwachsene langsamere mittlere und maximale Sakkadengeschwindigkeiten und kürzere Sakkadenamplituden sowie die Tendenz, weniger Sakkaden durchzuführen als die jüngeren Erwachsenen. Dies bestätigt Laborergebnisse, die zeigten, dass ältere Erwachsene ein engeres Sichtfeld haben. In der zweiten Studie wurde eine signifikante Beeinträchtigung von Sakkadengeschwindigkeit und -amplitude, hauptsächlich in der vertikalen Ebene, bei PSP-Patienten im Vergleich mit IPD-Patienten und gesunden Kontrollprobanden festgestellt, was wiederum konsistent mit Laborergebnissen ist. Interessanterweise war die Beeinträchtigung bei einer standardisierten Fixationsaufgabe, die wir zusätzlich zu der Messung in der realen Welt durchführten, bedeutend stärker. Dies deutet darauf hin, dass Patienten im wirklichen Leben die durch ihre Krankheit hervorgerufenen Defizite zum Teil kompensieren können. Insgesamt haben diese beiden Studien die mobile Augenbewegungsmessung in der realen Welt als eine schnelle und leicht anwendbare Forschungsmethode und als ein Diagnosewerkzeug bestätigt und weiter etabliert.

Wir konnten also einige Fragmente zu dem noch bruchstückhaften Wissen hinzufügen, wie Wahrnehmung von top-down Prozessen wie Aufmerksamkeit und Wertung beeinflusst wird, wie perzeptuelle Mehrdeutigkeit im Gehirn verarbeitet wird und wie sich visuelle Wahrnehmung mit dem Alter und mit verschiedenen Parkinsonsyndromen verändert. Die Ergebnisse stützen die Hypothese von Wahrnehmung als einem aktiven Prozess, der nicht nur reflexhaft von Reizen getrieben wird, sondern auch auf Erwartungen und Aufmerksamkeit beruht. Über die Grundlagenforschung hinaus sind unsere Ergebnisse ein erster Schritt in Richtung praktischer Anwendungen. Die Parametrisierung von gesundem Altern und die Differentialdiagnose zwischen verschiedenen Krankheiten sind also mögliche zukünftige Einsatzgebiete der mobilen Augenbewegungsmessung.

1 Introduction

1.1 Visual Perception

Any moment of our awake life we are confronted with a world that contains a vast number of objects and events, resulting in a huge amount of potentially useful information. However, much information is irrelevant for our present goals and behavior while some is absolutely essential. Furthermore, the metabolic cost of neuronal processing is high, which limits our information-processing capacity (Lennie, 2003). Thus, to be able to behave adequately in a complex environment, it is indispensable to extract and filter information that is relevant for our behavior and cognitive goals (Barlow, 1961; Einhäuser & König, 2010). The process of acquiring knowledge about our visual environment using information contained in the visible light reflected or emitted by our environment is what we refer to as visual perception (Palmer, 1999). The filtering and processing cascade that leads to perception starts with a physical stimulus. To be processed in the brain, the stimulus is transduced into electrical activity. This transduction takes place at the respective sensory organ, e.g. the eye when visual stimuli are concerned. Subsequently, the electrical activity passes through many areas of the brain, from low-level, very modality-specific areas like the primary visual cortex, right up to temporal, parietal and frontal regions of the brain, where more high-level functions are located and conscious perception of the stimulus can be established. It is important to note that this chain of events is not only a unidirectional flow of information from the sensory organ to more and more complex processing stages in the brain in a so-called feedforward way. Instead, feedback signals from higher cortical areas can enter processing at earlier stages and thus can strongly influence how the sensory signal is processed further (Sekuler & Blake, 1985). One striking property of this processing in the brain is that, in the neocortex, most connections are local and excitatory. Due to these recurrent connections the relatively small afferent signals can be amplified, restored and selected for further processing (Douglas & Martin, 2007) while very few long distance connections make it possible for brain regions to influence activity in other parts of the brain. In cognitive terms, the concept of feedforward and feedback signal-processing chains translate into stimulus-driven and expectation- or hypothesis driven perception, respectively (Palmer, 1999). Historically, feedback signals were not always regarded as being of high significance for the process of perception. Perception used to be considered as a unidirectional stimulus-driven process driven mainly

by the sensory input (Engel, Fries, & Singer, 2001). However, more than two decades ago, a paradigm-shift happened in favor of the notion that higher-level cortical functions do influence perception. Until now evidence has grown supporting this notion of an interplay of stimulus-driven and expectation- or task-driven processes that form our perceptual experience. Sensorimotor prediction, for instance, based on object properties and trajectories, is important for perceptual anticipation and adequate motor control (Land & McLeod, 2000; Nusseck, Lagarde, Bardy, Fleming, & Bülthoff, 2007). Furthermore, value can update our internal map according to (expected) benefits and (expected) costs, facilitating the processing of reward-related stimuli (Della Libera & Chelazzi, 2009; Hickey, Chelazzi, & Theeuwes, 2010). Hence, in most situations, our perceptual experience is composed of stimulus-driven information and our expectations based on prior experience and other higher cortical functions. Accordingly, our perception does not provide us with an exact imprint of the physical world. In fact, it is selective and continuously adapted to our expectations, tasks and goals and equips us with the information we need to behave optimally in a given situation. This emphasizes its role as an integral part of sensorimotor integration that does not only guide behavior but in turn is influenced by the generated actions itself (Einhäuser & König, 2010; Engel, Maye, Kurthen, & König, 2013).

1.1.1 Action and Perception

The importance of the interplay between action and perception is obvious in many situations of daily living like sports, driving a car, or in social interaction. The "common coding" theory formalizes the idea of mutual influences between action and perception. It claims that there are shared representations of action and perception (Prinz, 1997), which implies bidirectional influences between them. The transfer from visual perception to action is an everyday experience of sighted people and has been studied intensely. For instance, observational motor learning was shown to improve motor execution as well as action perception and the so-called mirror neurons have been proposed as the neurophysiological basis for this mechanism (Lago-Rodríguez, Cheeran, Koch, Hortobagy, & Fernandez-del-Olmo, 2014). However, the inverse influence of action on perception is not as obvious and only gained interest more recently (Schütz-Bosbach & Prinz, 2007). One of the first studies to show this action-perception transfer, which is predicted by the common coding theory, was done by Hecht, Vogt, and Prinz (2001). Participants had to execute a timed sinusoidal arm movement as a motor task and later judge the time ratios of sinusoidal bar movements displayed on a screen as a visual task. The authors found transfer effects from action to perception as well as from perception to action. In a study by Beets, Rösler, Henriques, Einhäuser, and Fiehler (2010), participants were presented with an ambiguous stimulus (cf. section 1.3) that could be perceived either rotating clockwise or counterclockwise and percep-

tion was reported by rotating a manipulandum in the respective direction, which lead to a stabilization of perception. However, when perception was indicated by turning the manipulandum in the opposed direction, perception destabilized. In contrast, when observers turned the manipulandum in a predefined direction and reported perception via button presses, no effect of the action could be found. This suggests that action must be percept-dependent to influence perception (Beets et al., 2010). This finding has been supported and extended by another study using a rivalrous stimulus where a video of the action itself was one of two binocular rivalry stimuli. Imitating the shown action while it was perceived stabilized perception while an unrelated movement or verbal description of the stimulus did not (Di Pace & Saracini, 2014). These findings imply that action can transfer to perception, but only when it is percept-dependent.

Taken together, perception does not only influence and transfer to action but there is also evidence for the reverse effect. This strengthens the point that the goal of perception is to support adequate behavior which in turn can influence future perception. This tight link between action and perception becomes evident also in neurophysiological studies that show an action-relatedness of neuronal firing rates in visual areas of the brain (Gallant, Connor, & van Essen, 1998; Mazer & Gallant, 2003). Hence, even though perception does not provide us with an exact representation of the environment, it can be considered optimal given the constraints of the output, namely, behavior. Interestingly, this notion allows for an alternative interpretation of the Jamesian bottleneck of attention (James 1890, see section 1.2). In this alternative view, attention is needed not mainly because processing resources are limited, but because readout and potential actions at a time are limited (Einhäuser & König, 2010), which emphasizes the importance of action and the mutual influences of perception and action even more.

1.1.2 Higher-level influences on perception

Since perception is composed of stimulus-driven "bottom-up" information and task- or goal-driven "top-down" influences, the question arises, as to how much each of the two factors contributes in certain situations. In this section, we will focus on the latter. The effect of top-down influences on perception becomes more prominent the more unclear the physical stimulus is. In an early study by Bruner and Potter (1964), blurred pictures of natural scenes were shown to participants. Blur was then reduced and participants had to report what they had been seeing. Initial blur and exposure time were varied and the greater the initial blur and the longer the exposure time was, the longer the recognition turned out to take. The authors interpreted this finding as the influence of the initial interpretation of the blurred picture, which, even though it was wrong in most cases, influenced perception until the object could be clearly identified. In contrast, Bar (2003) stated that based on images, even when reduced to low spatial frequencies (i.e.

blurred images), an "initial guess" can be made since the set of possible interpretations is limited. This confirms findings from Schyns and Oliva (1994) who found that early recognition is based on a coarse image scale while information are later refined by paying attention to the finer image scale. Hence, when blurred pictures are presented, top-down facilitation can take place. This is also in line with physiological findings that the fastest visual pathway ("magnocellular" pathway) transports low spatial frequency information (Maunsell, Nealey, & DePriest, 1990) and can thus lead to a quick recognition even of a blurred picture. These conflicting findings can potentially be reconciled considering the amount of blur and taking into account the reverse hierarchy theory. This theory states that the process of perception is top-down guided and only if these top-down information do not suffice to fulfill the task, information from lower visual areas are taken into account due to their better signal-to-noise ratio (Ahissar & Hochstein, 2004). In the study by Bruner and Potter (1964), this top-down influence led to a wrong interpretation of the stimuli because of the high amount of blur and this expectation was only changed when evidence for a different interpretation was clear based on the bottom-up information. The use of natural scenes instead of single objects made the task even more difficult in their study. In the second set of findings, only single objects and only slightly blurred pictures were considered. Thus, the pictures still yielded sufficient information to recognize the objects via top-down initial guesses and thus lead to fast recognition.

Apart from expectation and prior history, another candidate function to possibly alter perception is valuation or, more specific, expected outcome like reward or punishment. Expected outcome influences many of our decisions with higher expected rewards being typically preferred over low expected rewards or punishment, which has been studied intensely in the field of neuroeconomics (see Rangel, Camerer, and Montague, 2008 for review). Similarly, an influence of value on attention has been reported: In tasks requiring visual selective attention, the processing of reward-related features is facilitated (Della Libera & Chelazzi, 2009; Hickey et al., 2010), even several days after reward has been applied (Della Libera & Chelazzi, 2009) and even if deploying attention to the previously rewarded features was counterproductive for a given task (Hickey et al., 2010; see Figure 1.1). Furthermore, value affects attentional learning (Della Libera & Chelazzi, 2009; Della Libera, Perlato, & Chelazzi, 2011). Hence, given the direct influence of attention on perceptual appearance (Carrasco, Ling, and Read, 2004; see section 1.2), it is possible that value also modulates perception through attentional processes. Overlapping neuronal circuits for attention and reward were found in monkeys (Stănişor, van der Togt, Pennartz, & Roelfsema, 2013) and humans (Rothkirch, Schmack, Deserno, Darmohray, & Sterzer, 2014), rendering a link between attentional and reward-related mechanisms very likely. Alternative to the possibility of an indirect modulation of perception through selective attention, as suggested by the aforementioned studies, it is also conceivable that value exerts a direct influence on perception by upmodulating perceptual representations (Seitz,

Kim, & Watanabe, 2009). Up to now, however, no study has directly investigated whether value exerts a direct effect on perception.

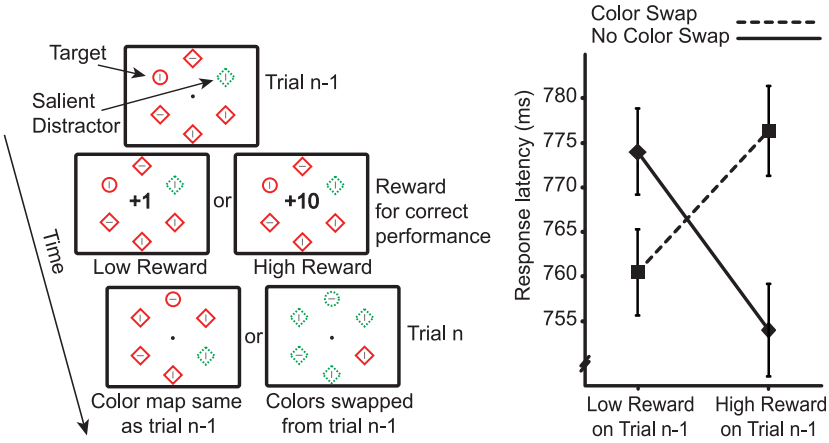

Figure 1.1: Effect of reward on attention. General paradigm and results from Hickey et al. (2010). Left panel: Observers had to report the orientation of the bar surrounded by the shape singleton. In 80 % of the trials a salient color singleton was presented as distractor. After each trial, either high (+10) or low (+1) reward was applied. Stimuli color and reward magnitude were varied between trials. Right panel: Reaction times for reporting the correct target as a function of reward magnitude in the previous trial and on whether color swapped between trials or remained the same. When there was no color swap, observers were fast after a high reward and slow followingn a low reward, whereas when there was a color swap, observers were slow following high reward. This indicates that the previously rewarded color strongly captured attention even though it indicated a distractor and thus had to be ignored. Even when reward predicted the likelihood of a color swap, reaction times were qualitatively similar.

1.2 Visual Attention

Already James (1890) pointed out one of the most important characteristics of attention: "it implies withdrawal from some things in order to deal effectively with others". So selective visual attention not only involves the process of assigning processing resources to one item, but also entails withdrawal of processing resources from others. Hence, it is a highly competitive mechanism allowing humans to prioritize aspects of their environment for efficient processing. The set of cognitive processes underlying these filtering mechanisms in cluttered visual scenes are called visual attention (McMains & Kastner, 2009). Even though attention can be shifted independently of eye movements by voluntarily holding fixation ("covert" visual attention), attentional orienting and gaze shifts are undoubtedly closely linked in everyday life ("overt" visual attention). For instance,

eye movements are preceded by a shift in attention to the position where the eyes will land (Deubel & Schneider, 1996). There is even evidence for shared neuronal circuitry of attention and eye movements (Rizzolatti, Riggio, Dascola, & Umiltá, 1987; Posner, 1980), which yields eye movements a good measure of attentional allocation in space.

It has long been agreed on that eye movements are guided by a so-called priority map which is composed of bottom-up factors like stimulus salience that are external to the observer and top-down influences that are internal to the observer (Desimone & Duncan, 1995; Posner, 1980; Kastner & Ungerleider, 2000; Zelinsky & Bisley, 2015). Top-down factors can be selection goals of the observer according to the relative task and selection history (Awh, Belopolsky, & Theeuwes, 2012). But selection history, particularly in the context of reward, can contradict the selection goals chosen according to a given task. For instance, Hickey et al. (2010) found that, in a reward paradigm, processing of previously reward-related features was facilitated such that attention was deployed there. This was even the case if participants knew that attending to reward-related features would be counterproductive for the task (see Figure 1.1). Hence, Awh et al. (2012) proposed to subdivide top-down processing into the two processes of goal-driven selection and selection history. Hence, according to the authors, the priority map that guides attention is constructed of bottom-up attentional control, goal-driven selection and selection history.

1.2.1 Attentional influences on perception

As can be derived from the strong interplay between eye movements and attention, attention also influences visual perception directly. In an influential paper, Carrasco et al. (2004) showed that attention alters the appearance of a grating by increasing its apparent contrast. In that study, covert visual attention was investigated, that is, the gratings were presented in the absence of eye movements in the periphery of the subject's visual field. When one grating was primed by a dot appearing at its future location, participants judged it as being of higher contrast than without priming. Another striking example of how attention influences perception is a phenomenon termed "inattentional blindness" (Mack & Rock, 1998). It refers to the failure of noticing clearly visible items when attention is deployed to a different task. Simons and Chabris nicely demonstrated this phenomenon using a movie of basketball players where participants should count the passes of one team. At some point a gorilla entered the scene for a few seconds, which was unnoticed by more than half of the participants (Simons & Chabris, 1999). Another example of attentional effects on perception is the attentional blink (AB), where perception's processing limitations in time become obvious (Raymond, Shapiro, & Arnell, 1992). This will be discussed in more detail in section 1.5.

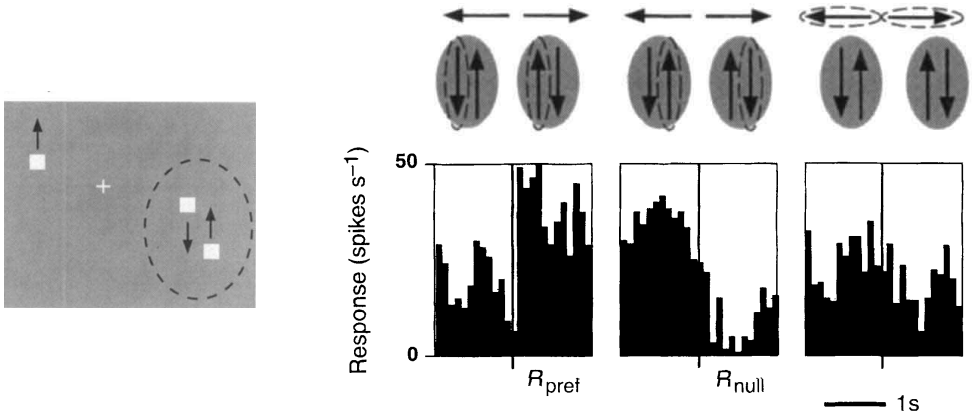

Figure 1.2: Neural basis of biased competition. Paradigm and responses of a neuron in MT. Left: The monkey had to fixate the small cross and covertly attend to one of the three moving dots, the respective dot was indicated at the beginning of each trial. The dashed oval corresponds to the RF of the neuron and the dots inside the RF where moving back and forth into opposite directions which corresponded to the neuron's preferred direction and the opposite direction, respectively. Right: Responses of the neuron showed clear dependence on attentional allocation. In the trials corresponding to the left and central histogram, a dot inside the RF was attended (marked by dashed oval), while in the right panel a dot outside the RF was attended. When a dot inside the receptive field was attended, the response varied strongly according to this dot's moving direction with a high firing rate when the dot moved in the preferred direction (here: upwards) and a low firing rate when the dot moved into the opposite direction. Importantly, the moving direction of the unattended dot in the RF did not influence the neuron's response. In contrast, the activity was relatively unmodulated when a dot outside the RF was attended (Treue & Maunsell, 1996).

1.2.2 Attention as biased competition

The competitive nature of attention and neurophysiological findings gave rise to the framework of "biased competition" (Desimone & Duncan, 1995), which is opposed to the "spotlight of attention" view inspired by James (1890) and its later extension to the "zoom-lens model" (Eriksen & James, 1986). In the zoom-lens model, processing resources are assigned to one location in the visual field of variable size while the rest is ignored. As this view is inherently spatial it cannot account for feature-based attention. In contrast, in the model of biased competition, attention is seen as an emergent property of many neural mechanisms that resolve competition by setting biases according to task demands in order to assign processing resources to the winning representation and to control action (Desimone & Duncan, 1995). This model complies with many physiological findings (Figure 1.2). In these experiments, two stimuli, one of which the cell is selectively responsive for, while for the other it is not, are brought into the same receptive field (RF). When the RF is not attended, the cell's response falls between the responses to the two single stimuli. When one of the stimuli is attended, the cell's response only corresponds to the attended stimulus, so it behaves

as if only this stimulus was present in the RF (Treue & Maunsell, 1996; Chelazzi, Duncan, Miller, & Desimone, 1998). Importantly, this finding is not restricted to stimuli that differ in location but can also be observed in feature-based attention (Treue & Martinez-Trujillo, 1999). These findings strongly support the view that competition is biased in favor of the attended stimulus. Another implication of these findings is that attention modulates neuronal responses in the same neural substrate where stimulus competition takes place (Dieter & Tadin, 2011).

The idea of attention as biased competition finds its correspondence in many attentional theories. In the theory of visual attention (TVA; Bundesen, 1990) and its neural implementation (NTVA; Bundesen, Habekost, and Kyllingsbæk, 2005), attention is described as a race between competing items for visual short-term memory. It is based on two mechanisms, namely, selection of objects (filtering) and selection of features (pigeonholing). Biases can be implemented on both levels. In a thorough review, Deco and Rolls (2005) derived a unifying theory that accounts for many electrophysiological and behavioral findings. In this framework cognitive behavior arises from attentional states held in short-term memory that by top-down processing influence visual processing using biased competition.

Taken together, the framework of biased competition formalizes the competitive nature of attention. Priority control is executed by setting biases to resolve competition in order to behave according to task demands. Using this framework, many experimental findings can be explained.

1.3 Rivalry

Perception of the world is not only complex, as pointed out in section 1.1, but also inherently ambiguous. The sensory systems provide under-constrained input about the external sensory sources and thus, congruent experiences can only

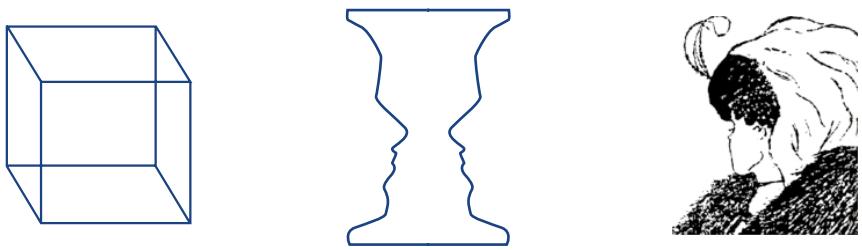

Figure 1.3: Ambiguous pictures. Each of these three pictures can be interpreted in at least two different ways: The Neckercube (Necker, 1832; left) can either have its front pointing to the lower right or to the upper left; the face-vase illusion (Rubin, 1915; center) can be perceived either as one vase or as two faces looking at each other; Boring's woman (Boring, 1930; right) can either be perceived as a young woman or as an old woman. Pictures from Blake (2001).

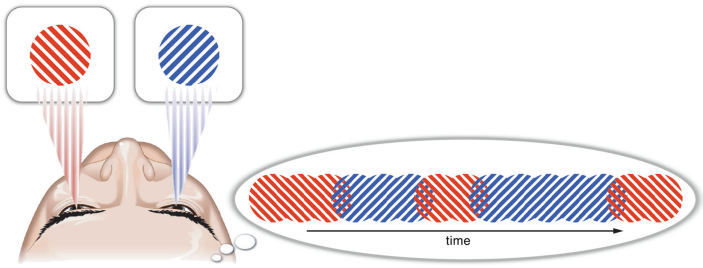

Figure 1.4: Binocular Rivalry. Binocular rivalry occurs when both eyes are presented with different stimuli. The observer perceives irregular perceptual alternations between the two stimuli with usually only one stimulus being perceived at a time (Dieter & Tadin, 2011).

be inferred from such fragmentary information. Helmholtz termed this process "unconscious inference" ("unbewusste Schlüsse", von Helmholtz, 1867). What we perceive is the most likely interpretation of the world, based on our prior knowledge and experience (Sekuler & Blake, 1985). If two interpretations are equally likely, this results in rivalry, a phenomenon where different interpretations of the world compete for awareness. Rivalry can be roughly subdivided into perceptual rivalry and binocular rivalry, which share some but not all properties (Klink, van Ee, & van Wezel, 2008; Meng & Tong, 2004). Prominent examples of perceptual rivalry are the Necker cube (Necker, 1832), Rubin's face-vase (Rubin, 1915), or Boring's woman (Boring, 1930) (Figure 1.3). In all these illusions, there are several possible interpretations of the drawings and perception might switch between them. The other form of visual ambiguity is binocular rivalry, a situation where both eyes are presented with different stimuli. This is typically implemented using a mirror stereoscope, first used by Wheatstone in 1838. The stimuli can be simple forms like gratings of different colors and orientations or more complex stimuli like a house and a face (Blake, 2001). It is critical that the stimuli are sufficiently distinct such that they cannot be merged into one percept. The perceptual outcome is an alternation between the two percepts (Figure 1.4). Disregarding short transition durations, binocular rivalry is typically exclusive at any given point in time, at least for small stimuli (Blake, O'Shea, & Mueller, 1992). For larger stimuli, sometimes fragments of both stimuli are seen at the same time, while at any location at any given point in time one stimuli exclusively dominates; this phenomenon is termed piecemealing. Measures of interest when studying binocular rivalry are the times and durations in which one percept dominates perception ("dominance durations"), the total amount of time one stimulus dominates in a trial ("dominance"), and the amount of alternations between percepts ("switches") throughout a trial. These measures follow well-defined rules, which can be condensed into three hallmarks of binocular rivalry.

1. *Heavy-tailed distribution.* Dominance durations typically follow a heavy-tailed ("leptokurtic") distribution (Levelt, 1967). A gamma-function has been proposed to fit these distributions (Levelt, 1967; Leopold & Logothetis, 1996), but the exact shape is still debated (Brascamp, van Ee, Pestman, & van den Berg, 2005).

2. *Levelt's propositions.* A set of four rules, proposed by Levelt (1968), describes the relation of stimulus strength to dominance and mean dominance duration. Stimulus strength can refer to different stimulus properties such as contrast and luminance. Some of the rules have been challenged recently (Moreno-Bote, Shpiro, Rinzel, & Rubin, 2010; Brascamp, van Ee, Noest, Jacobs, & van den Berg, 2006; Bossink, Stalmeier, & de Weert, 1993) and the revised form of the rules is:
   i. Increase of the stimulus strength in one eye will increase the predominance of the stimulus (Levelt, 1968).
   ii. Changes in stimulus strength of one eye affect the mean dominance duration of the highest contrast eye (Brascamp et al., 2006).
   iii. Alternation frequency is maximal at and symmetric around equi-dominance (Moreno-Bote et al., 2010).
   iv. Increase of the stimulus strengths in both eyes will increase the alternation frequency (Levelt, 1968).

3. *Survival probability after blanking.* When the stimulus periodically disappears (is blanked) for a certain duration and the duration is sufficiently long, perception stabilizes, even for a substantial period of time (Orbach, Ehrlich, & Helen, 1963; Leopold, Wilke, Maier, & Logothetis, 2002). This effect is reversed into destabilization when the blanking durations are short (Orbach et al., 1963). That is, the probability that the percept reappears after blanking ("survival probability") increases with increasing blanking duration.

1.3.1 Higher-level influences on rivalry

Since perception can be affected by higher-level cortical processes, also rivalry can be influenced by attention, expectation and prior knowledge. It was shown that prior exposure to an unambiguous version of a rivalry stimulus influences perception of the ambiguous stimulus in perceptual rivalry. Leeper (1935) showed that in Boring's woman, when a less ambiguous version of the drawing was shown before the ambiguous version was presented, there was a tendency to interpret the drawing in the way of the unambiguous version seen beforehand (Figure 1.5). In binocular rivalry, attention also largely influences perception, which is to some

Figure 1.5: Example for the effect of priming on perception. When disambiguated versions of the illusion are shown before the ambiguous version is presented, observers most likely perceive the previously seen interpretation of the illusion. Left: disambiguated version of the young woman; center: ambiguous version; right: disambiguated version of the old woman. (Sekuler & Blake, 1985)

degree well in line with the concept of biased competition where attention resolves competing representations according to task demands (Dieter & Tadin, 2011). General attention to the stimulus speeds up alternations (Paffen, Alais, & Verstraten, 2006) while attention to one specific stimulus increases its perceptual dominance and decreases switch rate (Ooi & He, 1999; Meng & Tong, 2004; van Ee, van Dam, & Brouwer, 2005). Helmholtz believed that he could keep one pattern dominant as long as he wished by exerting attentional control in simple stimuli and in more complex stimuli by performing a task like counting (von Helmholtz, 1867). However, later studies showed that even though participants could influence their perception during rivalry, full control over the dominant percept is impossible (Breese, 1899; Meng and Tong, 2004; for review see Paffen and Alais, 2011). Thus, the arising competition in rivalry cannot be resolved entirely. Understanding rivalry in the framework of biased competition (cf. section 1.2.2) also allows more insight into how it could be influenced by attention. In biased competition, the attentional bias resolves competition within the same neural substrate as stimulus competition appears, which implies that if competition is limited to low-level bottom-up mechanisms, options for attentional modulations are limited (Dieter & Tadin, 2011). Thus, attention should have a higher impact on perceptual rivalry, where high-level interpretations of the stimulus compete, as compared to binocular rivalry, where stimuli presented to the two eyes compete, because this competition could be resolved at very early stages of visual processing. This notion is supported by two studies (Meng & Tong, 2004; van Ee et al., 2005) who found a much stronger attentional modulation on perceptual rivalry than on binocular rivalry stimuli.

Taken together, rivalry is a good model to study perception since the physical stimulus that is presented remains constant while subjective perception changes. This makes rivalry a particularly well-suited tool for studying the influence of internal higher-level processes like attention, motivation or emotion on perception.

Furthermore, its tight link to attentional processes and the potential common framework of biased competition make it a valuable tool to particularly study the interplay between attention and visual competition.

1.4 Modeling Rivalry

Binocular rivalry, through its alternating nature, provides a good tool to study the underlying competitive neural processes in the brain that lead to perceptual alternations. To gain deeper insight into how properties of perceptual alternations originate from neural activity and network connectivity, it is essential to relate rivalry characteristics to networks in the brain. There has been substantial progress in the field of modeling rivalry and many models perform well in replicating some characteristics of rivalry. Most models are based on mutual inhibition and adaptation, which is either implemented on a spiking-neuron level (Laing & Chow, 2002; Wilson, 2003, 2005) or in a firing-rate model (Lago-Fernández & Deco, 2002; Freeman, 2005; Noest, van Ee, Nijs, & van Wezel, 2007). Several of these low-level accounts of rivalry share the property of multiple representational levels, accounting for the different processing stages of the brain (Lago-Fernández & Deco, 2002; Freeman, 2005). Moreno-Bote, Rinzel, and Rubin (2007) introduced a noise-driven attractor-based network that can be implemented in firing-rate and spiking-neuron networks and can be extended to multistability phenomena. All of these models produce heavy-tailed dominance durations and some account for particular statements of Levelt's propositions. Only one model covers blanking (Noest et al., 2007), when it is refined with a multi-timescale extension (Brascamp, Pearson, Blake, & van den Berg, 2009). But even though the high number of models present by now reflects the importance and the interest in a model of rivalry, a physiologically plausible and comprehensive model accounting for all rivalry characteristics mentioned earlier is still lacking.

Due to the exclusivity of binocular rivalry, at least at a representational level accessible to an observer's introspection and visual awareness, perception is unique and interpretations do not intermix. As long as one interpretation is dominant the others are suppressed. In neural terms this implies that as long as a particular interpretation is dominant ("wins"), the other interpretations are less active ("lose"). The ongoing dynamics of this competition can be conceptualized as a winner-take-all (WTA) process where the winner changes as a function of time (Yuille & Geiger, 2003). These WTA models are frequently used to model attentional processes (Itti & Koch, 2000; Lee, Itti, Koch, & Braun, 1999; Hahnloser, Douglas, Mahowald, & Hepp, 1999). Often in these models, recurrent connections between units are employed to select the "winning" percept (Hahnloser et al., 1999; Rutishauser & Douglas, 2009), thus picking up a basic property of how the brain processes information in the neocortex (Douglas & Martin, 2007).

There is convincing evidence supporting a link between attentional mechanisms and rivalry. In rivalry and selective attention, stimulus selection is accomplished by assigning more processing resources to one stimulus at the expense of others. Even though this effect is more dramatic in rivalry, this suggests that the mechanisms of the two phenomena are closely linked (Leopold & Logothetis, 1999). Also the concept of attention as biased competition links it directly to rivalry, where several visual interpretations compete for awareness. Not only the competitive nature relates the two processes, but also the attentional bias that influences perception during rivalry (Paffen & Alais, 2011) and thus to a considerable degree helps resolving competition. Taken together, due to the similarities between rivalry and selective attention in behavioral measures and underlying processes, using an attentional model for modeling rivalry seems very appropriate.

At the circuit level, WTA-type behavior emerges if a population of excitatory neurons is recurrently connected to itself and shares a common inhibitory signal, which leads to excitatory neurons competing with each other (Yuille & Geiger, 2003; Rutishauser, Douglas, & Slotine, 2011). The statistics of connectivity between cortical neurons indicate that recurrent connectivity between cortical neurons is a fundamental feature of organization of the neocortex (Binzegger, Douglas, & Martin, 2004; Douglas & Martin, 2004; Douglas, Koch, Mahowald, Martin, & Suarez, 1995). Such recurrent connections provide the excitatory and inhibitory feedback necessary to give rise to WTA-type dynamics. These dynamics are of interest for rivalry because they combine a discrete binary winner while remaining sensitive to the continuously changing inputs for all possible interpretations including the current winner as well as all current losers. This continued sensitivity to activity in all neurons allows the network to change its winner continuously. coupled WTA circuits can implement states that remain stable in the absence of external input. They thus have memory of previous inputs, making their behavior dependent both on the current input as well as on previous input as represented by the current state of the network. Thus, WTA networks are promising for modeling rivalry.

1.5 Rapid Visual Processing

Another way to challenge perception, apart from presenting ambiguous stimuli, is to limit the time in which stimuli are available for processing. In the according experiments, pictures are flashed only very briefly (20-200ms) and can be followed by a mask to make the task more difficult. This method of rapid image presentation can be extended by showing a series of pictures - including targets and distractors - in direct succession, a so-called rapid serial visual presentation (RSVP). These experiments testing rapid scene processing frequently employ detection, categorization or delayed match-to-sample tasks (Thorpe, Fize, & Marlot, 1996; Evans

& Treisman, 2005; Yao & Einhäuser, 2008; Gegenfurtner & Rieger, 2000).

The processing of briefly presented natural stimuli can be remarkably quick (Potter & Levy, 1969). This was also demonstrated by Thorpe et al. (1996) in a rapid go/no go categorization task where pictures were presented in isolation. In natural-scene pictures, animals could be detected with near-ceiling performance when presented for only 20ms (unmasked). Category-specific event-related potentials (ERPs) in the EEG could be observed as early as 150ms after stimulus onset. In another paradigm, observers were presented with two images, one in each hemifield, and had to respond with a saccade towards the hemifield where they had seen an animal. Saccadic reaction times could be as short as 120ms (Kirchner & Thorpe, 2006). The authors concluded that there must be a very fast route linking visual processing to the programing of saccades. This rapid processing is not restricted to animal detection but is also valid for objects like vehicles (van Rullen & Thorpe, 2001) and for detecting human faces under animal faces (Rousselet, Macé, & Fabre-Thorpe, 2003). Surprisingly, in tasks requiring speeded responses, there seems to be a robust global minimal reaction time (minRT, 250-290ms), which is defined as the first time bin where correct responses start to significantly outnumber incorrect responses (Fabre-Thorpe, 2011). This minRT is valid across different experiments, irrespective of the target stimuli (Fabre-Thorpe, 2011) and even does not change after intensive training (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001). These findings raise the questions what the attentional influences are in this context and which features are relevant for rapid visual processing.

For testing the attentional influences on rapid visual processing, one particularly interesting phenomenon called the attentional blink (AB; Raymond et al., 1992) can be employed. It can be studied in RSVP experiments with more than one target and is observed when two of the targets appear in close temporal succession: Perception of the second target (T2) is impaired when it is shown between 200 and 700ms after a first target (T1), so it is less frequently detected than without T1. If T2 immediately follows T1, this decreased detection rate is usually absent (Raymond et al., 1992). In past experiments, mainly artificial items like numbers and letters have been used to study the AB (Raymond et al., 1992; Chun & Potter, 1995) but in more recent studies, also natural-scene pictures were employed (Evans & Treisman, 2005; Einhäuser, Koch, & Makeig, 2007). Measuring the detection or categorization performance over time makes the AB a good indicator of the temporal cost of the selective allocation of attention (MacLean & Arnell, 2012).

It has been argued that the rapid processing of natural scenes can be nearly attention-free. In a rapid natural-scene categorization task, embedded in a dual-task paradigm, participants could identify animals and vehicles from distractors in the near absence of attention (Li, van Rullen, Koch, & Perona, 2002). This was not the case for differentiating T's from L's or other tasks using artificial stimuli. The authors concluded that high-level representations can be accessed without the deployment of attention. Another study, in turn, questioned this

result. Evans and Treisman (2005) conducted a series of RSVP experiments including one or more targets. In their AB experiments they found that target category (animal/vehicle) was essential for the AB effect: If the targets had to be identified and belonged to different categories, the AB was more severe than when two targets from the same category were presented. Furthermore, when both targets only had to be detected, the AB vanished for same-category stimuli and was only marginally present for different-category stimuli. The authors came to the conclusion that detection of animals and vehicles in natural scenes is mediated by rapid feature analysis while attention-demanding binding is necessary for identification and localization (Evans & Treisman, 2005).

To answer the question, which features are relevant for rapid visual processing of naturalistic stimuli, several studies varied different stimulus features in natural scenes. Wichmann, Drewes, Rosas, and Gegenfurtner (2010) tested whether the power spectrum of images has an impact on animal detection, which has been suggested by Torralba and Oliva (2003). Performance was essentially independent of the power spectrum, but classification was facilitated by the spectral cue without being caused by it. In an animal/no animal categorization task using different presentation durations, fast mechanisms relied on shape while the integration of shape and texture was somewhat slower but made detection more robust. Color and luminance played virtually no role in this task (Elder & Velisavljević, 2009). In an RSVP detection task, removing color also had no influence on performance (Meng & Potter, 2008), while in a task using rapid presentation without postmask, Delorme, Richard, and Fabre-Thorpe (2010) found a small but significant influence of color on accuracy in an animal categorization task for late (>325ms) responses while not for short ones. In this study, the most crucial features for accurate and quick detection were a typical animal posture and the area occupied by the animal. In a rapid animal/no animal categorization task, Wichmann, Braun, and Gegenfurtner (2006) found only a 2-3% increase in performance for color stimuli as compared to grayscale stimuli. In an RSVP task, Yao and Einhäuser (2008) found also little effect of color on detection of animals in natural scenes.

Even though color plays only a minor role in the detection and categorization of natural-scene stimuli, there is a frequently reported effect of color on memory. Interestingly, in the aforementioned study by Yao and Einhäuser (2008), there was an effect of color on recognition memory and confidence, so participants were more confident when they reported seeing a colored animal than a grayscale animal. In addition, when both, a grayscale and a color picture of an animal, were presented in a picture stream, the colored animal was more likely to be reported. This suggests that color has an effect on retrieval from memory rather than an effect on detection. Besides being relevant for retrieval, color is also advantageous for encoding into memory (Gegenfurtner & Rieger, 2000). Hence, color seems to have little influence on rapid detection in natural scenes. However, the association of the function of color vision with attentional processes (Frey, Honey, & König, 2008; Maunsell & Treue, 2006; Motter, 1994) and encoding into memory

(Gegenfurtner & Rieger, 2000) raises the question if and how color influences the AB.

1.6 Eye Movements

Eye movements are an overt manifestation of attentional orienting and thus provide a good indication of where attention is allocated under natural conditions (cf. section 1.2). Due to the distribution of photo receptors on the human retina, which shows a high density only in a small spot ("fovea"), humans are forced to redirect gaze frequently to see a large part of their environment in high resolution. Since the locations the eyes can target in a certain amount of time are limited, relevant target locations have to be selected while irrelevant ones have to be filtered out. These locations are determined by endogenous or exogenous attention. Hence, visual attention allows us to select the information that is most relevant to ongoing behavior (Chun & Wolfe, 2005). Since the eyes are guided by the attentional value of the environment, eye movements can be used as a measure to assess relevant, attention-drawing locations in visual scenes. The locations where the eyes rest (fixation locations) indicate regions of interest while the fast ballistic eye movements (saccades) between fixations indicate where to fixate next and thus what part of the scene will be investigated in more detail next. If a moving target is followed by the eyes, a tracking eye movement called smooth pursuit can be observed. Eye movements proved to be a good indicator of attentional allocation and are also part of the filtering system to reduce complexity in a manifold environment.

Eye-movement patterns are highly task-dependent. This was first investigated by Buswell (1935), who in a large study investigated how different people look at art pictures given different tasks. He already used an elaborate method to track the eyes and found that the viewing pattern of participants highly differed between a search condition - looking for a person in the window of a skyscraper - as compared to carefully studying the same picture without a task (Figure 1.6). Yarbus (1967), one of the pioneers in modern eye-tracking, confirmed this finding by investigating eye-movement patterns while participants looked at a complex picture (Repin's "Unexpected Visitor"). He asked observers different questions about the picture and found that eye-movement patterns highly differed between trials where different questions were asked. This finding confirms that eye movements are driven by cognitive processes and reflect the current task.

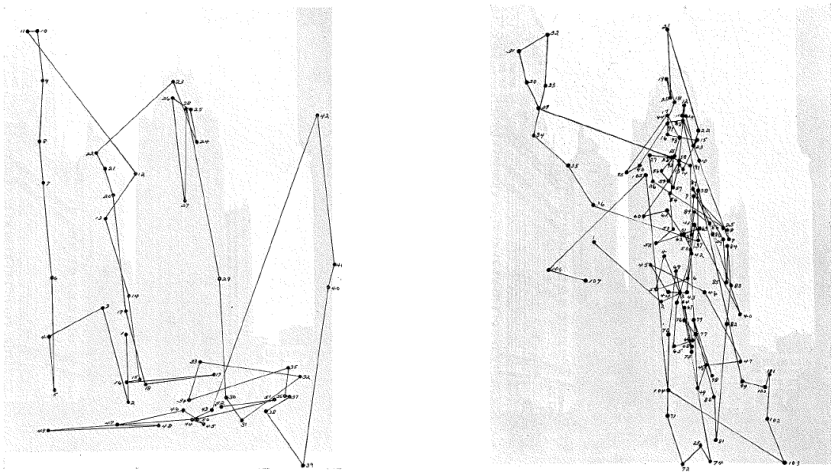

Figure 1.6: Eye-movement patterns demonstrating task-dependency of eye movements. Recordings from one subject looking at a picture of the Tribune Tower in Chicago, which is located in the center of the picture. Numbered dots indicate fixations, lines mark gaze shifts and the original picture is outlined in the background. The left pattern was obtained giving no explicit instruction while in the session where the eye-movement pattern on the right was recorded, the observer had to look for a person in the skyscraper's window (Buswell, 1935).

1.6.1 Eye movements in the real world

Most eye-movement phenomena have been studied using artificial stimuli or videos in laboratory settings. This is a good approach to study very specific aspects of vision, but for understanding how vision copes with a realistic natural environment and how it relates to action, more natural stimuli and settings are required (Einhäuser & König, 2010). The advent of mobile eye trackers made it possible to study eye movements in real-life situations including sports (Hayhoe, Mennie, Sullivan, & Gorgos, 2002; Land & McLeod, 2000; Fairchild, Johnson, Babcock, & Pelz, 2001), everyday activities like food or tea preparation (Land & Hayhoe, 2001; Land, Mennie, & Rusted, 1999) and driving (Land & Lee, 1994; Land & Tatler, 2001). In addition to investigating eye movements in real-life situations, mobile eye tracking provides a means of studying the interplay between action and perception in natural situations. Eye-head coordination as well as the interplay of body movements and eye movements can be studied in everyday tasks (Land et al., 1999) or free exploration (Einhäuser, Schumann, et al., 2007). During tasks, nearly all eye movements are directed towards task-relevant objects or locations and the roles of fixations could be roughly divided into *locating* an object, *directing* the effector (in most cases the hand) to a new object or location, *guiding* the relative move of several objects and *checking* the state of some external variable (Land et al., 1999; Land & Hayhoe, 2001). During natural exploration without a specific task, eye movements mostly occur along the cardinal axes and are highly dependent on the environment (Einhäuser, Schumann, et al., 2007). They

co-occur with head movements more often than would be predicted by an independence assumption. Most co-occurring eye and head movements point in opposite directions, indicating the gaze-stabilizing role of eye movements, but a substantial fraction also points in the same direction (Einhäuser, Schumann, et al., 2007), serving larger gaze shifts. Motivated by these findings, one study compared eye-movement patterns recorded in real life to those recorded in laboratory settings using videos recorded by a head-centered camera attached to the eye-tracking device. They showed that oculomotor behavior significantly differs between the different conditions ('t Hart et al., 2009). Thus, the method of mobile eye tracking has to be used to further investigate how laboratory findings translate into real life.

1.6.2 Effects of aging on eye movements

With increasing age, the function of the human visual system declines. Aging goes along with poorer visual acuity (Klein, Klein, Linton, & De Mets, 1991) as well as an impaired perception (e.g., perceived heading direction, Lich and Bremmer, 2014; motion perception, Billino, Bremmer, and Gegenfurtner, 2008). Most visual performance measures such as reading, face recognition, low contrast vision and attentional visual field area were shown to decline progressively with age, starting in the sixth decade (Haegerstrom-Portnoy, 2005). In addition to these perceptual parameters, also eye-movement parameters are changing with age. Saccadic latencies are increased (Munoz, Broughton, Goldring, & Armstrong, 1998; Moschner & Baloh, 1994) while smooth-pursuit gain decreases (Moschner & Baloh, 1994). Some studies find a decrease in saccade peak-velocity with age (Irving, Steinbach, Lillakas, Babu, & Hutchings, 2006; Moschner & Baloh, 1994) while others don't find an effect of age on saccade peak-velocity (Munoz et al., 1998). In several visual tasks like visual search and singleton-detection, older adults relied relatively more on top-down processes like expectations and prior knowledge than on low-level features (Whiting, Madden, Pierce, & Allen, 2005; Madden, 2007), which could partly be due to their decline in visual performance. Acik, Sarwary, Schultze-Kraft, Onat, and König (2010) investigated the effects of age on viewing natural scenes in an experiment where natural and artificial complex scenes were shown before a later patch-recognition task had to be performed. Participants were children (7-9 years), young adults (19-27 years) and older adults (>72 years). The authors found a U-shaped performance curve with young adults outperforming the other groups and a progressively decreasing feature-dependence with age. Thus they confirmed the age-related shift from bottom-up to top-down reliance in visual tasks. Interestingly, children that used more top-down information and older adults that showed a higher feature dependence than their peers performed better in the task, showing a benefit of higher explorativeness with age.

Thus, the study contributed to the understanding of developmental changes in natural and task-dependent viewing.

1.6.3 Application of eye-movement analysis as diagnostic tool

Since eye movements are not only easily accessible but also reliable and quick to measure, they have gained in importance as a tool for detecting functional impairments in the brain in the last decades (Leigh & Zee, 2006; Leigh & Kennard, 2004). Particularly saccades are valuable tools for the clinical neurosciences since they are amongst the best-understood movements and follow a very stereotypical trajectory which makes them highly comparable. The underlying neuronal mechanisms that are involved in planning, generating and executing different kinds of saccades are largely known, making them a good indicator of neuronal diseases affecting these brain areas (Leigh & Kennard, 2004). Clinical studies using eye tracking are also interesting for basic research since, in well-controlled studies, the differences in eye-movement parameters can be traced back to the neurological differences between subject groups. Thus, the independent variable in these experiments is the disease the patients suffer from and it can be varied by including different patient groups. This makes patient studies a valuable method for understanding how the respective disease and the related underlying neurological dysfunctions influence eye movements.

To make eye-tracking techniques broadly available, it is important to make them readily applicable and easy to handle for physicians. Since most eye-tracking systems are stationary and not easily transportable, the use of mobile eye trackers would make clinical application more effective and convenient for both, the physicians and the patients. Additionally, the range of potential applications would increase by making it possible to study body-, head- and eye-movements at the same time, which is particularly interesting for the study of movement disorders.

1.7 Studies

This thesis consists of five studies that investigate visual perception. In study 1, we test the effect of value and attention on perception using binocular rivalry. In the second study, we model binocular rivalry using a WTA network that is frequently used for modeling attention. The third study employs an RSVP task to investigate the effect of color on rapid scene processing. The final two studies test how mobile eye-tracking can be applied to parametrize healthy aging and to differentiate between two neurodegenerative diseases (PSP and IPD).

2 Studies

Study 1:
Reward modulates perception in binocular rivalry
Marx S. & Einhäuser W.
Journal of Vision, 2015

Study 2:
Competition with and without priority control:
Linking rivalry to attention through winner-take-all networks with memory
Marx S., Gruenhage G., Walper D., Rutishauser U. & Einhäuser W.
Annals of the New York Academy of Sciences, 2015

Study 3:
Rapid serial processing of natural scenes:
Color modulates detection but neither recognition nor the attentional blink
Marx S., Hansen-Goos O., Thrun M. & Einhäuser W.
Journal of Vision, 2014

Study 4:
Effects of aging on eye movements in the real world
Dowiasch S., Marx S., Einhäuser W. & Bremmer F.
Frontiers in Human Neuroscience, 2015

Study 5:
Validation of mobile eye-tracking as novel and eficient means
for differentiating progressive supranuclear palsy from Parkinson's disease
Marx S., Respondek G., Stamelou M., Dowiasch S., Stoll J., Bremmer F., Oertel W.H., Höglinger G.U. & Einhäuser W.
Frontiers in Behavioral Neuroscience, 2012

Journal of Vision (2015) 15(1):11, 1–13 http://www.journalofvision.org/content/15/1/11 1

Reward modulates perception in binocular rivalry

Svenja Marx Neurophysics, Philipps-University, Marburg, Germany

Wolfgang Einhäuser Neurophysics, Philipps-University, Marburg, Germany

Our perception does not provide us with an exact imprint of the outside world, but is continuously adapted to our internal expectations, task sets, and behavioral goals. Although effects of reward—or value in general—on perception therefore seem likely, how valuation modulates perception and how such modulation relates to attention is largely unknown. We probed effects of reward on perception by using a binocular-rivalry paradigm. Distinct gratings drifting in opposite directions were presented to each observer's eyes. To objectify their subjective perceptual experience, the optokinetic nystagmus was used as measure of current perceptual dominance. In a first experiment, one of the percepts was either rewarded or attended. We found that reward and attention similarly biased perception. In a second experiment, observers performed an attentionally demanding task either on the rewarded stimulus, the other stimulus, or both. We found that—on top of an attentional effect on perception—at each level of attentional load, reward still modulated perception by increasing the dominance of the rewarded percept. Similarly, penalizing one percept increased dominance of the other at each level of attentional load. In turn, rewarding—and similarly nonpunishing—a percept yielded performance benefits that are typically associated with selective attention. In conclusion, our data show that value modulates perception in a similar way as the volitional deployment of attention, even though the relative effect of value is largely unaffected by an attention task.

Introduction

In decision making, the role of expected outcome (i.e., reward or punishment) and related variables has been studied intensely (Deco, Rolls, Albantakis, & Romo, 2013; Glimcher & Rustichini, 2004; Preuschoff, Bossaerts, & Quartz, 2006) and led to increasingly sophisticated theories of motivational learning (Rescorla & Wagner, 1972; Watkins & Dayan, 1992). Similarly, the interest in the (neuro-) physiological foundations of reward processing, which dates back to Pavlov (1927) and Skinner (1938), has flourished since dopamine's role as signal for reward prediction error was unveiled (Schultz, Apicella, Scarnati, & Ljungberg, 1992), eventually leading to neuroeconomics (Glimcher & Rustichini, 2004) as a new field within the neurosciences. Especially in nonhuman primates, decision-making experiments frequently employ perceptual tasks, such as discovering coherent motion in random dot patterns (Newsome & Pare, 1988). Despite the use of perceptual tasks, surprisingly little research has addressed direct effects of reward and punishment on perception per se. This is even more remarkable, given that contemporary models of perception under ambiguity and decision-making under uncertainty often use the same "Bayesian" formalism (Bülthoff & Yuille, 1996; Freeman, 1994; Kersten, Mamassian, & Yuille, 2004). At least in the context of perceptual rivalry, pupillometric data suggests shared neural mechanisms for decision making under uncertainty and the resolution of perceptual ambiguity (Einhäuser, Stout, Koch, & Carter, 2008). When viewing natural perception as inferring a unique perceptual interpretation from underconstrained sensory information (Von Helmholtz, 1867), it is tempting to think of perception as a decision process among the infinite number of possible interpretations. Under this hypothesis, valuation processes that modulate cognitive decision making should similarly exert a direct influence on the perceptual interpretation of constant stimuli.

Valuation processes can in principle exert an influence on perception in two ways: First, valuation may modify perception through selective attention, which directly alters perceptual appearance (Carrasco, Ling, & Read, 2004). Indeed, there is mounting evidence for effects of reward on attention: The processing of reward-associated features is facilitated in tasks requiring visual selective attention (Della Libera & Chelazzi, 2009; Hickey, Chelazzi, & Theeuwes, 2010) and rewards affect attentional learning (Della Libera & Chelazzi, 2009). The effects of reward on attentional

Citation: Marx, S., & Einhäuser, W. (2015). Reward modulates perception in binocular rivalry. *Journal of Vision*, *15*(1):11, 1–13, http://www.journalofvision.org/content/15/1/11, doi:10.1167/15.1.11.

doi: 10.1167/15.1.11 Received September 3, 2014; published January 14, 2015 ISSN 1534-7362

processing are present even several days later (Della Libera & Chelazzi, 2009), and differ between different forms of automatic orienting such as space- and object-based attention (J. Lee & Shomstein, 2013). Second, rewards may exert a direct influence on perception by modulating perceptual representations without or in addition to attentional mechanisms (Seitz, Kim, & Watanabe, 2009).

Binocular rivalry, a situation where the two eyes are presented with dissimilar stimuli, is an ideal paradigm to test direct effects on perception, since the stimulus remains unchanged while perception alternates between two alternatives (percepts). Effects of attention on rivalry are well established: Attention speeds up the alternations between percepts (Paffen, Alais, & Verstraten, 2006), and attention to one stimulus increases its perceptual dominance (Ooi & He, 1999; Van Ee, van Dam, & Brouwer, 2005). Whether reward exerts a similar effect on perceptual dominance and whether it uses attentional mechanisms or acts in addition to attention, is unknown.

In the present study we used binocular rivalry to test the hypothesis that reward has a direct effect on perception. In a first experiment (Experiment 1), we tested whether explicitly rewarding one percept has a similar effect as attending it. In separate parts we either asked observers to attend one of the percepts or explicitly associated a reward with seeing one of the percepts. To circumvent the issue of relying on observers' report, we used the optokinetic nystagmus (OKN) to objectively measure at any point in time which percept observers were subjectively experiencing (cf. Fahle, Stemmler, & Spang, 2011; Naber, Frässle, & Einhäuser, 2011). In a second experiment (Experiment 2), we again explicitly rewarded one percept, but in addition asked observers to perform an attentionally demanding task either on the rewarded percept, the nonrewarded percept, or both, thus generating three attentional conditions. In a separate part of Experiment 2, reward was replaced by punishment, with otherwise unchanged instructions or stimuli. This allowed us to test the hypothesis that reward and punishment bias perception even when attention is engaged in a different task.

Methods

Participants

Eight participants (seven female, 27.0 ± 3.66 years) participated in Experiment 1. This number was decided upon prior to the experiment based on estimates derived from previous rivalry studies (e.g., Naber et al., 2011), which showed that—despite considerable interindividual variability in absolute dominance durations—behavioral effects are typically robust, in that their qualitative direction (sign) can be expected to be consistent across individuals. Eight participants (six female 25.5 ± 2.98 years) participated in Experiment 2, with one participant (#6) participating in both experiments (15 participants in total). The number of participants in Experiment 2 was chosen to match Experiment 1, and was decided upon after the conclusion of Experiment 1, and prior to starting Experiment 2. Before the experiment, participants gave written informed consent. All procedures were in accordance with the Declaration of Helsinki and approved by the local ethics committee (Ethikkommission FB04).

Setup and stimuli

Stimuli were generated using Matlab (Mathworks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and Eyelink toolbox (Cornelissen, Peters, & Palmer, 2002) extensions. They were displayed on two 21-inch Syncmaster CRT screens (Samsung, Seoul, South Korea), each set to 1280 × 1024 pixels spatial and 85 Hz temporal resolution, and presented dichoptically at a viewing distance of 30 cm by using a mirror stereoscope. Each observer's left eye position was tracked by a noninvasive eye-tracking device (EyeLink 2000, SR Research, Osgoode, ON, Canada) at 500 Hz. The eye tracker's infrared camera and illuminator were positioned behind the mirrors, which were transparent to infrared light (cold mirrors), such that the eye-tracking setup was not visible to the observer.

In both experiments, sine-wave gratings with a spatial frequency of 0.21 cycles per degree were presented to both eyes for 180 s in each trial. Gratings were of different color (red/green), differently oriented (± 20°) and drifted upward perpendicular to their orientation at a speed of 14.25°/s (Figure 1a). The gratings were presented in a circular aperture with a diameter of 30° in Experiment 1 and of 21° in Experiment 2. The aperture was surrounded in both eyes by the same blue annulus that could vary in width from 0° to 3.4° (60 pixels).

In reward trials, the annulus width grew proportionally to the amount of reward, in punishment trials proportionally to the monetary punishment. In trials without reward or punishment, the width remained constant at a value between 0° and 3.4° that was chosen randomly for each trial.

OKN slow phase as measure of perceptual dominance, assignment of reward

While most binocular-rivalry experiments in humans rely on the observers' subjective reports regarding their

Journal of Vision (2015) 15(1):11, 1–13

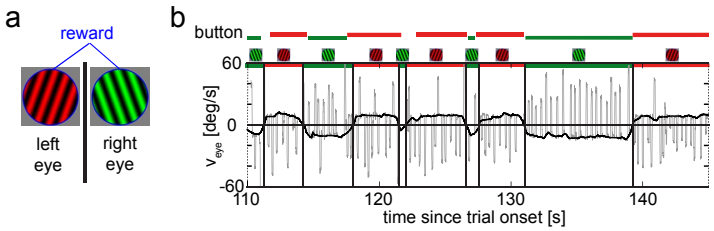

Figure 1. Stimuli and objective measure of perceptual state. (a) Binocular rivalry stimulus: Using a mirror stereoscope, each eye was presented a different drifting grating; the gratings differed in color, orientation, and motion direction to robustly induce binocular rivalry; the blue annulus signaling reward (in some conditions of Experiment 2: punishment) was presented identically to both eyes. (b) Example excerpt eye-trace of one observer (Observer #4); gray trace: raw velocity; black trace: interpolated OKN slow phases; black vertical lines: time of zero-crossings of OKN slow phase that define a perceptual switch; red/green bars inside of graph: perceptual state as defined by OKN; red/green bars on top of graph: button corresponding to red/green grating pressed (gaps imply no button pressed at the respective time point). In general button presses, which are only available in the active-report condition, aligned well with the switch times inferred from the OKN slow phase, which were available in all conditions and on which all analysis was based.

perceptual experience, this is suboptimal in the present context for at least two reasons: First, observers could strategically choose to report their percept nonveridically to maximize reward (i.e., they could "cheat"); second, the requirement to report one's percept could interfere with attentional tasks. Since the direction of OKN's slow phase closely replicates the observer's perceptual experience when viewing drifting gratings (e.g., Naber et al., 2011), we used this objective measure throughout. For applying reward, eye velocity was calculated online by differentiation of the eye's raw horizontal position collected at a frequency of 85 Hz. If three successive samples were in the required direction and velocity range (between 3.6°/s and 30°/s), reward (or punishment) was increased by one point.

For calculating dominance durations and alternation rates, the eye traces were processed offline. First, horizontal eye velocity was obtained by differentiation of raw horizontal eye position. Then all OKN fast phases were removed by applying thresholds to convolution-filtered (square-smoothing window of 0.1-s width) eye-velocity traces (>15°/s) and acceleration (>100°/s^2; Figure 1b). All removed parts of the velocity trace were then interpolated using a piecewise cubic Hermite interpolation. An objectively measured switch from one percept to the other was then defined as a zero crossing of the resulting horizontal OKN slow phase. Dominance durations were defined as the time between successive switches; alternation rate was defined as the number of switches per time. For the active-report conditions of Experiment 1, we also verified the correspondence between button presses and OKN-defined dominance phases and found them to be well matched, with the exception of short dominance durations being missed by the observers' subjective reports (Figure 1b; see also Naber et al., 2011).

Procedure

Assignment of reward and punishment

In reward/punishment trials, the online OKN analysis resulted in a reward/punishment point for each sample in which the rewarded/punished percept was dominant (provided the velocity criterion was met for at least three successive samples; see above). Each aggregated 180 points resulted in a one pixel increase of the blue annulus. This increase was sufficiently smooth to look continuous to the observers. We allowed a maximum annulus of 60 pixels, which corresponded to 10,800 points or 127 s (10,800/85 Hz) of dominance (70.5% of the trial) of the respective percept. In Experiment 1, the maximum reward (60 pixels) corresponded to 1€ of actual money with linear mapping of points to Euros. In the reward blocks of Experiment 2, 60 pixels corresponded 0.5€; in the punishment blocks of Experiment 2, 0 pixels corresponded to 0.5€ and 60 pixels to 0€.

Experiment 1

Experiment 1 consisted of three different conditions. In the reward condition, participants were instructed before trial onset that they were going to be rewarded for seeing one of both colors. Reward was then indicated by the width of the blue annulus surrounding the drifting gratings. In the attention condition, before trial onset participants were instructed to attend to one of both colors. In the no-instruction condition, stimuli were presented without specific instructions regarding reward or attention. In half of the trials of each condition, participants were in addition instructed to report the grating's drifting direction by pressing and holding one of two buttons (active-report condition); in

the other half they were just passively viewing (passive-viewing condition). Including the active-report condition in Experiment 1 allowed us to probe possible interactions of the requirement to report with reward and to verify the OKN analysis (see above). For comparability between conditions, however, all analysis in both conditions was based on dominance durations as inferred from the OKN data.

The experiment was split in four sessions of three blocks each. Each block consisted of four trials. The instruction condition (reward, attention, none) was constant in each block, but the assignment of reward/attention to color and the report condition changed between trials within blocks.

Experiment 2

To test whether the effect of rivalry on reward prevailed when attention was engaged either on the rewarded stimulus or elsewhere, in Experiment 2 we aimed at increasing the attentional load on the observers. Unlike typical dual-task situations, the eyes in our paradigm were in constant motion, following the perceived grating. The equivalent to performing a task at fixation is therefore to perform a task that is spatially locked to the grating. Consequently, we asked observers to perform a task with the drifting grating. Specifically, participants were instructed to detect a change in duty cycle of the grating, which lasted for three frames (35 ms) and occurred 30 times per trial and grating in random intervals of 1 to 6 s. Participants were instructed for which grating they had to report changes; they reported their detection by a button press. Observers could either be instructed to report changes only in one grating (full attention) and ignore the other (attention away) or to report changes in both gratings (split attention). Reward and punishment instructions were given in addition to the attentional instructions. This yielded a 2×3 design: Besides being rewarded or not, a stimulus could receive full attention (duty cycle change only to be monitored for this stimulus), split attention (both stimuli monitored), or attention away (other stimulus monitored). For the time a grating was dominant, we calculated performance as the fraction of duty-cycle changes that an observer reported within 1 s, divided by the total number of duty-cycle changes that occurred in the respective grating during its dominance. For the correctly reported duty-cycle changes in each grating, we in addition computed the average reaction time from the onset of a duty-cycle change to its report. To ensure task compliance, participants were also told that they were only given the money if performance in the detection task "was sufficiently good."

To avoid interference with the attention task, observers were not required to report their percept; that is, all conditions of Experiment 2 in this respect corresponded to the passive viewing condition of Experiment 1. Other than the duty-cycle changes, which themselves did not induce changes in dominance, and a slight reduction in size (see above), to ease monitoring the reward-signaling annulus during the attentional task, stimuli were identical to Experiment 1.

Experiment 2 consisted of two types of blocks. In the reward blocks, participants were rewarded for seeing one of the two possible percepts, indicated by the blue annulus. In the punishment blocks, participants were penalized for seeing one of the two percepts. The penalty was also indicated by the blue annulus with a thicker ring, meaning less money. Before every trial, participants were informed about which was the rewarded/penalized stimulus and in which grating they had to execute the detection task. This resulted in three conditions per block: The rewarded/penalized percept could equal the percept in which the detection task had to be executed, the rewarded/penalized percept could be the percept that should not be attended for the task, and the rewarded/penalized percept could be one of the two attended percepts.

The conditions were randomized in blocks consisting of six trials each and every condition appeared four times, resulting in 12 trials per reward and punishment condition. In one experimental session, one reward and one punishment block each consisting of six trials was measured and each participant took part in two experimental sessions.

Statistical analysis

For comparisons between two conditions paired t tests were used (within-subject design), and for comparisons involving more than one factor or more than two levels per factor, repeated measures ANOVAs were used, treating observers as repeated measures. As measures of effect size, Cohen's d is reported for t tests and partial eta square (η_p^2) is reported for ANOVAs. All statistical analysis was conducted using Matlab.

Results

In Experiment 1, we tested the effect of reward on perceptual dominance in binocular rivalry, and separately the effect of attention. Using the OKN as an objective measure allowed us to include conditions in which observers actively monitored and reported their current percept and those in which they just passively viewed the stimulus.

Journal of Vision (2015) 15(1):11, 1–13

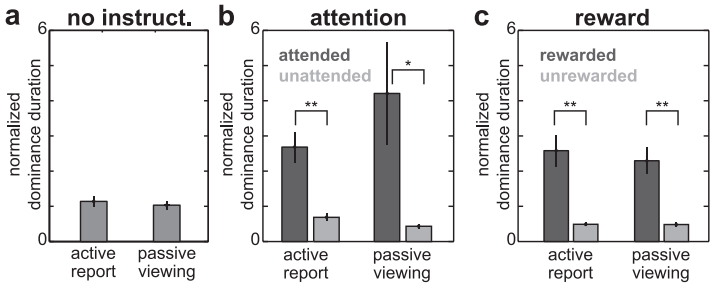

Figure 2. Normalized dominance durations of Experiment 1. (a) Perceptual dominance durations when no instruction regarding attention or reward was given, left: active report; right: passive viewing. (b) Dominance durations when observers were instructed to attend one of the gratings (dark gray: attended grating; light gray: other grating). (c) Dominance durations when reward was provided proportional to viewing duration of one grating (dark gray: rewarded grating; light gray: other grating). In all graphs, bars denote *M*, error bars *SEM* over $N = 8$ observers. All dominance durations were normalized across all conditions within each observer, panels (a–c). Raw dominance durations of each individual are given in the Appendix. Significance markers refer to paired posthoc tests (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

Robust induction of rivalry

To induce OKN reliably, we deliberately used large gratings as stimuli. Large stimuli in rivalry frequently result in mixed percepts (piecemealing). Although we did not query piecemealing explicitly, we had instructed participants during active-report conditions in Experiment 1 to report exclusive dominance. Under this instruction, the times in which no percept was reported (or both percepts were reported simultaneously) give an indication how frequently the percept was unclear; that is, times of possible piecemealing. These periods only accounted for 7.15% ± 3.58% ($M \pm SD$ over observers) of active-report trials, and we observed no difference between the attention and the reward condition [attention: 9.67% ± 11.5%, reward: 4.20% ± 3.42%, attention vs. reward: $t(7) = 1.19$, $p = 0.27$, $d = 0.086$]. Even though we cannot exclude piecemealing in full, more than 90% of time observers reported an exclusive, unambiguous percept.

Reward and attention similarly bias perception

When observers received no instructions regarding attention or reward, they had dominance durations between 0.578 and 1.50 s ($M \pm SD$ over observers: 0.978 ± 0.364 s), and there was no difference between active reporting and passive viewing, $t(7) = 1.36$, $p = 0.22$, $d = 0.48$, paired t test. There was no strong bias for either percept, neither in the active-report (56.2% ± 9.3% green percept dominant) nor in the passive-viewing (54.1% ± 11.0%) condition. Due to the high interobserver variability in absolute dominance durations, which is typical for rivalry, for the remainder we normalized dominance durations in each observer by dividing all data per observer by the median dominance duration over the whole experiment. The relative effects were, however, qualitatively consistent across observers (all individual data are shown in the Appendix). With the normalized dominance durations, we still observed no significant difference between active report and passive viewing, $t(7) = 1.04$, $p = 0.33$, $d = 0.37$ (Figure 2a).

When instructing observers to attend one of the stimuli, its dominance duration increased significantly as compared to the unattended stimulus irrespective of whether the dominance was actively reported or not [2 × 2 repeated-measures ANOVA; main effect attended vs. unattended: $F(1, 7) = 8.62$, $p = 0.022$, $\eta_p^2 = 0.55$; main effect active report vs. passive viewing: $F(1, 7) = 1.40$, $p = 0.28$, $\eta_p^2 = 0.17$; interaction attention × report: $F(1, 7) = 2.99$, $p = 0.127$, $\eta_p^2 = 0.30$; Figure 2b]. When observers were instructed that one stimulus was rewarded, the respective stimulus similarly became significantly more dominant [main effect rewarded vs. unrewarded: $F(1, 7) = 33.48$, $p = 0.0007$, $\eta_p^2 = 0.83$] irrespective of passive or active viewing [main effect: $F(1, 7) = 0.27$, $p = 0.62$, $\eta_p^2 = 0.037$; interaction: $F(1, 7) = 0.29$, $p = 0.61$, $\eta_p^2 = 0.040$; Figure 2c]. To compare the attention and reward sessions directly, we in addition performed a three-way ANOVA on all data with factors VIEWING (active, passive), INSTRUCTION_TYPE (reward, attention), and AFFECTED_PERCEPT (rewarded/attended vs. other). As expected, there was a main effect of whether the percept was affected by instruction [i.e., the rewarded or attended percept as compared to the respective other percept: $F(1, 7) = 13.65$, $p = 0.008$, $\eta_p^2 = 0.66$], but no effect of INSTRUCTION_TYPE [$F(1, 7) = 2.90$, $p = 0.13$, $\eta_p^2 = 0.29$] or VIEWING [$F(1, 7) = 0.63$, $p = 0.45$, $\eta_p^2 = 0.083$] and no interactions (all Fs < 3.28, all ps > 0.11). Hence, reward and attention to a percept both yielded a significant increase in its dominance duration, which—

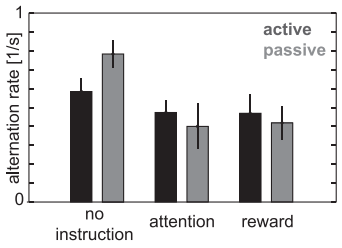

Figure 3. Alternation rates in Experiment 1. From left to right: Trials with no specific instruction, trials with attended stimulus, trials with rewarded stimulus. Dark gray: active-response condition; light gray: passive-viewing condition.

when applied independently—was indistinguishable in magnitude. Hence Experiment 1 demonstrates an effect of reward on perception, but leaves open whether this effect is achieved merely by the observers allocating attention to the rewarded stimulus.

As alternative measure to dominance durations, we calculated alternation rates, the number of switches in perception per unit time (Figure 3). A 3 × 2 repeated measures ANOVA revealed a significant main effect of INSTRUCTION_TYPE [no instruction, attention, or reward; $F(2, 14) = 8.30$, $p = 0.0042$, $\eta_p^2 = 0.54$], no main effect of VIEWING [active or passive; $F(1, 7) = 1.28$, $p = 0.29$, $\eta_p^2 = 0.15$] and a significant interaction between INSTRUCTION_TYPE and VIEWING [$F(2, 14) = 5.70$, $p = 0.016$, $\eta_p^2 = 1$]. Posthoc paired t tests showed that the effect of instruction type resulted from a difference between no instruction and the other two conditions (ts > 2.58, ps < 0.037, exception: active: attention vs. no instruction; Table 1), which did not differ from each other (ts < 0.43, ps > 0.68). Hence, the results on alternation rates show the same pattern as the analysis of dominance durations: Reward modulates rivalry to a similar extent as attention.

Reward modulates perception at constant attentional load

In Experiment 2, we tested the interaction between attention and reward. As an attentional task, observers had to respond to changes in the duty cycle of the attended grating or gratings (Figure 4a) while one of the two percepts was rewarded. Observers complied with this attentional instruction, and responded within 1 s to duty-cycle changes for full-attention stimuli in 77.7% ± 6.8%, for split-attention stimuli in 82.3% ± 3.1%, and (incorrectly) for attention-away stimuli only in 7.0 ± 6.7% of cases. We found a main effect of attention, $F(2, 14) = 73.89$, $p = 3.6 \times 10^{-8}$, $\eta_p^2 = 0.91$) and of reward, $F(1, 7) = 35.06$, $p = 0.0006$, $\eta_p^2 = 0.83$, on dominance durations. Although there was a significant interaction, $F(2, 14) = 5.63$, $p = 0.016$, $\eta_p^2 = 0.45$, posthoc tests showed that for each level of attention (full, spilt, away), reward had a significant effect on dominance durations (Figure 4b). This pattern was consistent across individuals (see Appendix). Reward had an effect on dominance durations in every attentional condition [full: $t(7) = 3.85$, $p = 0.0063$, $d = 1.36$; away: $t(7) = 9.07$, $p = 4.0 \times 10^{-5}$, $d = 3.21$; split: $t(7) = 3.76$, $p = 0.0071$, $d = 1.33$]. Comparing the same reward condition between different attentional conditions revealed significant effects of attention on rewarded stimuli [full vs. away: $t(7) = 6.39$, $p = 0.00037$, $d = 2.26$; full vs. split: $t(7) = 4.29$, $p = 0.0036$, $d = 1.52$; away vs. split: $t(7) = 4.02$, $p = 0.0050$, $d = 1.42$], and on unrewarded stimuli [full vs. away: $t(7) = 9.83$, $p = 2.4 \times 10^{-5}$, $d = 3.47$; full vs. split: $t(7) = 7.20$, $p = 0.00018$, $d = 2.54$; away vs. split: $t(7) = 8.91$, $p = 4.6 \times 10^{-5}$, $d = 3.15$]. This demonstrates that even at the same instruction regarding attention, dominance durations are in addition modulated by reward.

	Conditions	*t* value	*p* value	Cohen's *d*
Experiment 1				
Active	No instruction vs. attention	$t(7) = 2.58$	0.037	0.91
	No instruction vs. reward	$t(7) = 1.73$	0.13	0.61
	Attention vs. reward	$t(7) = 0.11$	0.91	0.04
Passive	No instruction vs. attention	$t(7) = 2.74$	0.029	0.97
	No instruction vs. reward	$t(7) = 3.40$	0.011	1.20
	Attention vs. reward	$t(7) = 0.42$	0.69	0.15
Experiment 2				
Rewarded	Attended vs. unattended	$t(7) = 3.80$	0.0067	1.34
	Attended vs. split	$t(7) = 3.57$	0.0091	1.26
	Unattended vs. split	$t(7) = 2.25$	0.060	0.79
Unpunished	Attended vs. unattended	$t(7) = 4.36$	0.0033	1.54
	Attended vs. split	$t(7) = 4.11$	0.0045	1.45
	Unattended vs. split	$t(7) = 0.062$	0.95	0.022

Table 1. Statistical measures and effect sizes of posthoc comparisons of alternation rates.

Journal of Vision (2015) 15(1):11, 1–13

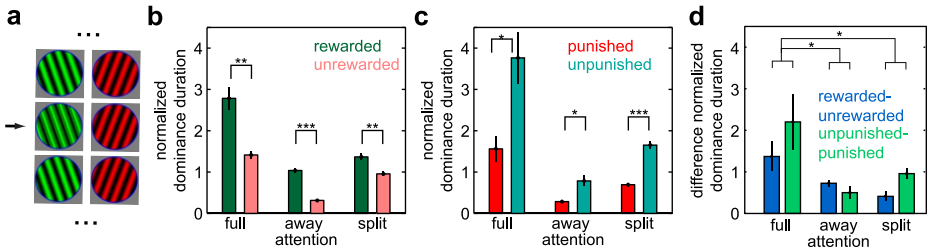

Figure 4. Attention task and normalized dominance durations of Experiment 2. (a) Illustration of attentional task (time running from top to bottom). Every 1 to 6 s one of the gratings changed its duty cycle for 35 ms (change in green grating indicated by arrow); observers had to report these changes either in one of the gratings (attention full) and not in the other (attention away) or in both (attention split). In addition one of the gratings was rewarded/punished. (b) Dominance duration for rewarded (green) and nonrewarded grating (light red), split by attentional conditions. (c) Dominance duration for punished (red) and nonpunished grating (light blue), split by attentional conditions. (d) Differences of dominance durations between rewarded and unrewarded (blue) and between unpunished and punished condition (green). In all graphs bars denote *M*, error bars *SEM* over $N = 8$ observers. Dominance durations were normalized within each observer across the whole Experiment 2. For raw dominance durations of each individual see Appendix. Significance markers in (b) and (c) refer to paired posthoc tests in each attention condition. To avoid crowding the figures, only significance markers for the contrasts between reward/punishment and no reward/no punishment are depicted; for details on other effects, see text. Significance markers in (d) refer to the posthoc test for effects between pairs of attentional levels (see text for details). The difference between away and split is not significant ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

Absence of punishment biases perception similarly to presence of reward

In a separate part of Experiment 2, we replaced the reward instruction by an instruction regarding monetary punishment. Stimuli were exactly identical to the reward part, and the instructions only differed in so far that the increasing blue annulus now signaled a reduction in monetary gain. Again observers complied well with the attentional instruction and responded faithfully to duty-cycle changes in attended stimuli only (full: 75.3% ± 3.4%; split: 78.8% ± 6.9%; away: 2.1% ± 2.5%). Qualitatively, the nonpunished percept behaved similarly to the rewarded percept in the other experimental part, and vice versa (Figure 4c; all individual data are shown in the Appendix). Indeed, there was a significant main effect of attention, $F(2, 14) = 20.80$, $p = 6.4\text{x}10^{-5}$, $\eta_p^2 = 0.75$, and a significant main effect of punishment, $F(1, 7) = 20.12$, $p = 0.003$, $\eta_p^2 = 0.74$. Although we found an interaction, $F(2, 14) = 5.94$, $p = 0.014$, $\eta_p^2 = 0.46$, posthoc tests revealed significant differences for punishment versus no punishment at each attentional level [Figure 4c; full: $t(7) = 3.30$, $p = 0.013$, $d = 1.17$; away: $t(7) = 3.34$, $p = 0.012$, $d = 1.18$; split: $t(7) = 7.45$, $p = 0.00014$, $d = 2.63$]. Comparing the same punishment conditions between attentional conditions revealed effects of attention on the punished stimulus [full vs. away: $t(7) = 3.81$, $p = 0.0066$, $d = 1.35$; full vs. split: $t(7) = 2.67$, $p = 0.032$, $d = 0.94$; away vs. split: $t(7) = 6.63$, $p = 0.00030$, $d = 2.34$] and the unpunished stimulus [full vs. away: $t(7) = 4.66$, $p = 0.0023$, $d = 1.65$; full vs. split: $t(7) = 3.28$, $p = 0.014$, $d = 1.16$; away vs. split: $t(7) = 4.69$, $p = 0.0022$, $d = 1.66$]. In sum, punishing a percept had a similar effect to rewarding the competing percept: The more valuable percept increased in dominance for each attentional condition.

Direct comparison of reward and punishment

To compare the effects of reward and punishment directly, the difference between dominance durations of rewarded and unrewarded stimuli at the same attentional level was calculated (Figure 4d). A 3 × 2 ANOVA revealed a main effect of attentional level, $F(2, 14) = 9.69$, $p = 0.0023$, $\eta_p^2 = 0.58$, but no effect of instruction (reward/punishment), $F(1, 7) = 2.57$, $p = 0.15$, $\eta_p^2 = 0.27$, and no interaction, $F(2, 14) = 1.79$, $p = 0.20$, $\eta_p^2 = 0.20$. For the reward condition, posthoc paired t tests revealed significant differences between the full and split, $t(7) = 2.57$, $p = 0.037$, $d = 0.91$, as well as for the away and split, $t(7) = 3.19$, $p = 0.015$, $d = 1.13$, conditions while the difference between full and away only tended to be significant, $t(7) = 1.98$, $p = 0.088$, $d = 0.70$. In the punished condition, only the difference between the full and away condition were significant, $t(7) = 2.98$, $p = 0.021$, $d = 1.05$, while the other differences only tended to be significant [full vs. split: $t(7) = 1.93$, $p = 0.096$, $d = 0.68$; away vs. split: $t(7) = 2.31$, $p = 0.054$, $d = 0.82$]. The missing effects of instruction type and interaction indicate that the effect of reward and punishment are symmetric (rewarding one percept is equivalent to punishing) and both are similarly modulated by the attention task.

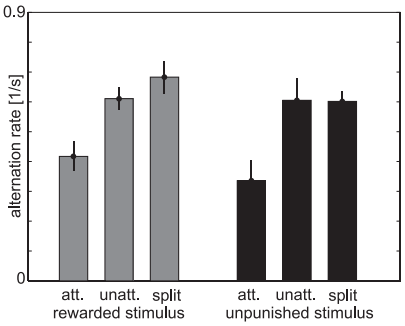

Figure 5. Alternation rates in Experiment 2. Left: sessions with reward, right: sessions with punishment. From left to right: trials with the rewarded stimulus attended, trials with the unrewarded stimulus attended, trials with split attention between rewarded and unrewarded stimulus, trials with the unpunished stimulus attended, trials with the unpunished stimulus unattended (i.e., the punished stimulus is attended), trials with split attention between punished and unpunished stimulus. Note that in contrast to Figure 4, data are here sorted by trial type, not by stimulus.

As for Experiment 1, we considered alternation rates as an alternative measure to dominance durations also for Experiment 2 (Figure 5). Since the alternation rate provides one number per trial, we distinguished three types of reward trials: The rewarded stimulus was attended (rew/att), one stimulus was attended and the other rewarded (rew/unatt), both stimuli were attended and one was rewarded (rew/split). To treat punishment analogous to the absence of reward, we sorted the three types of punishment trials according to the "unpunished" stimulus: Either the unpunished stimulus was attended (unp/att), the punished stimulus was attended and the unpunished was not (unp/unatt), or attention was split (unp/split). If punishment is indeed equivalent to the absence of reward, we expect no effect of instruction (reward vs. punishment) and no interaction between instruction and attention. If, however, the presence of reward or punishment has an arousing effect on the rewarded or punished stimulus, effects of reward absence and punishment should differ and modulate the attentional effect. Performing a 2×3 ANOVA with the factors of attention (att, unatt, split) and instruction (rew/unp) showed a main effect of attention, $F(2, 14) = 18.27$, $p = 0.00013$, $\eta_p^2 = 0.72$, but no effect of instruction, $F(1, 7) = 2.17$, $p = 0.18$, $\eta_p^2 = 0.24$, and no interaction, $F(2, 14) = 0.83$, $p = 0.46$, $\eta_p^2 = 0.11$. Posthoc t tests showed that for the rewarded and the unpunished stimuli, this effect of attention was entirely due to the difference between the attended condition on the one hand and the other two attentional conditions on the other hand (all ts > 3.56, all ps < 0.0092; Table 1), while there was no difference between the split and the unattended condition (ts < 2.25, ps > 0.059). This lack of effect for instruction supports the notion that for our paradigm, reward and absence of punishment have the equivalent effect, rendering the possibility that reward and punishment merely augment the rewarded/punished stimulus in an unspecific way unlikely.

Reward has similar effects on performance as attention itself

Value exerts an effect on perceptual dominance on top of the effect induced by the instruction to volitionally attend a percept to use it for a task. Since Experiment 1 had already suggested that the qualitative impact of reward and attention are similar, we here ask whether reward and punishment can also modulate performance for an attended percept akin to attentional effects. In the split-attention condition of Experiment 2, both the rewarded and the nonrewarded percept were not only present at the same time, but also received identical instructions regarding attention. Hence this condition allowed us to test whether the analogy between attention and reward at constant attentional load extended to typical attentional measures like performance or reaction time. Of the duty-cycle changes happening while the rewarded percept was dominant, 84.9% ± 3.5% were detected. During its dominance, the nonrewarded stimulus tended to be detected less frequently, 78.2% ± 7.6%, $t(7) = 2.28$, $p = 0.056$, $d = 0.81$. Conversely, in the punishment part of the experiment, the punished percept tended to be detected less frequently than the nonpunished percept, though this difference did not reach significance, 73.1% ± 16.3% vs. 80.0% ± 7.7%, $t(7) = 1.11$, $p = 0.30$, $d = 0.39$. A slight difference was also evident for reaction times, with reactions trending to be faster for rewarded stimuli, 429 ± 49 ms vs. 443 ± 61 ms, $t(7) = 2.12$, $p = 0.071$, $d = 0.75$, and being significantly faster for nonpunished stimuli, 440 ± 63 ms vs. 422 ± 65 ms, $t(7) = 2.38$, $p = 0.049$, $d = 0.84$, respectively. Even though the effects of reward and punishment on processing the already attended stimulus are comparably weak, their trend is in line with the notion that value has a similar effect on perception as attention.

Discussion

We exploited an objective measure of perceptual dominance in binocular rivalry to assess effects of reward and punishment on perception. Effects of reward were qualitatively similar to effects of attention. Nonpunishing showed similar effects as reward, sug-

gesting that the observed effects are specific to positive value and not a mere consequence of general stimulus relevance. Directing attention to one stimulus did not abolish effects of value (i.e., of reward or punishment). Irrespective of whether or not valuation and attention share the same mechanism, this suggests that volitional deployment of attention according to task demands leaves room for additional modulation of perception by value.

Our data show many commonalities between volitionally attending a stimulus and being rewarded for perceiving it. Both attention and positive value (i.e., reward or nonpunishment) up-modulate a stimulus' dominance and show a trend to improve processing (performance and reaction times) for a stimulus that is already attended. The latter may have three different, not mutually exclusive, explanations: First, the positively valued percept might have better visibility; second, positive value may improve processing directly; and third, positive value allocates additional attentional resources to an already attended stimulus. Provided our findings that reward can act in addition to volitionally deployed attention, the notion that reward effects are attentional in nature would be in line with the observation that reward can guide—or even capture—attention irrespective of other attention-guiding factors (Failing & Theeuwes, 2014).

Using an objective measure of perceptual dominance is critical for two reasons: First, since OKN cannot be controlled volitionally, we rule out that observers' report (rather than their perception) is biased by reward; that is, we avoid any possibility of "cheating." Second, when imposing an attentional task, we avoid that this interferes with the necessity of attending the rivalry stimulus for report. While other objective measures could be employed to assess an observers' perceptual state (e.g., fMRI decoding; Tong, Nakayama, Vaughan, & Kanwisher, 1998), the OKN has been established as robust measure of perceptual dominance in rivalry in humans (Fahle et al., 2011; Frässle, Sommer, Jansen, Naber, & Einhäuser, 2014; Naber et al., 2011) and animals (Fries, Roelfsema, Engel, König, & Singer, 1997; Logothetis & Schall, 1990).

It has long been known that attention influences perception during binocular rivalry. Endogenously attending to one percept can stabilize it, although perception cannot be controlled entirely (Breese, 1899; Meng & Tong, 2004; Ooi & He, 1999; van Ee et al., 2005). This stabilizing effect of voluntary control is stronger for perceptual than for binocular rivalry (Meng & Tong, 2004; van Ee et al., 2005). Attention modulates perception in binocular rivalry in different ways; exogenously cueing attention to the rivaling stimuli can initiate switches in rivalry and thus speeds alternations (Ooi & He, 1999; Paffen et al., 2006; Paffen & van der Stigchel, 2010) and, accordingly, drawing attention away from the rivaling stimuli slows alternation rate (Paffen et al., 2006). While perceptual rivalry can occur even without allocating attention to the rivaling stimuli (Pastukhov & Braun, 2007), binocular rivalry requires visual attention (Brascamp & Blake, 2012; Zhang, Jamison, Engel, He, & He, 2011). Using fMRI, S.-H. Lee, Blake, and Heeger (2007) showed that drawing attention away from the rivaling stimulus left rivalry-related activity in primary visual cortex (V1) but not in exstrastriate visual areas (V2, V3) where activity was rivalry-related when the stimuli were attended. Both our experiments confirm the basic effects of attention on rivalry. In turn, reward influenced dominance in rivalry as well as typical measures of attention, such as reaction times. This possibly points to a deeper conceptual link between visual selective attention and rivalry, which both are competitive processes, for which one stimulus is selected at the expense of others (Leopold & Logothetis, 1999).

Viewed in isolation, the finding that reward decreases overall alternation rates and increases dominance of the rewarded stimulus could be interpreted as perception being biased by stimulus relevance (Alpers, Ruhleder, Walz, Mühlberger, & Pauli, 2005) in a rather unspecific manner akin to arousal. Such a generic relevance effect would predict some up-modulation for the punished stimulus, making the effect on the unpunished stimulus in punishment trials at least smaller (if not reversed) than on the rewarded stimulus in reward trials. The fact that we instead find symmetry between presence of reward and absence of punishment suggests that the effect of reward is specific and not explained by stimulus relevance per se. Hence, the effect of valuation on perception is selective. In this selectivity, value is similar to attention (rather than to arousal).

Since reward and attention in our Experiment 1 had similar effects on perceptual dominance, and attention modulates perceptual appearance (Carrasco et al., 2004), it is tempting to assume that reward and punishment act through attentional processes. By engaging attention either elsewhere or on the rewarded stimulus, Experiment 2 tested whether rewards can modulate perception irrespective of other attentional demands. While this attentional task follows the logic of dual-task paradigms (e.g., D. K. Lee, Koch, & Braun, 1999; Pastukhov, Fischer, & Braun, 2009), there are several conceptual differences. First, there is no actual primary task, since rivalry in Experiment 2 is restricted to passive viewing; second, the secondary task is not conducted at fixation, but aligned with the movement of the stimulus. The former can be considered uncritical, provided that Experiment 1 showed no difference between active and passive conditions with respect to attentional and reward effects. The latter is necessary to keep the task fixed in

retinal coordinates, since the OKN stabilizes the dominant percept relative to the retina; thus the situation is similar to dual tasks at fixation for static stimuli. As with typical dual-task paradigms, however, there is no guarantee that residual attentional resources are made fully unavailable. Indeed, the trend to worse performance for unrewarded percepts may even be interpreted as evidence for value acting through attentional mechanisms: Value interferes with the attentional task and therefore the mechanisms mediating valuation and attention are not fully independent. In sum, dominance durations in Experiment 2 demonstrate that value modulates perception at any level of attention, implying that this volitional deployment of attention required for task performance leaves room for modulation of perception by value. In turn, the performance data suggest that attention and valuation share common mechanisms to bias perception.

Besides plenty of behavioral links between reward and attention (Della Libera & Chelazzi, 2009; Hickey et al., 2010; J. Lee & Shomstein, 2013), the interaction of attention and reward has been widely studied in the neurosciences. In nonhuman primates, even in V1, neurons' activities were predicted by reward value of the respective stimulus and neurons that exhibited strong value effects also showed strong attentional effects implying overlapping neuronal selection mechanisms for value and top-down attention (Stănişor, van der Togt, Pennartz, & Roelfsema, 2013). In humans, reward-associated distractors in a saccade tasks had large influence on saccade curvature even when they should be ignored, resembling behavior in tasks where distractors were made more salient (Hickey & van Zoest, 2012). This analogy to an increase in salience was also found in an electroencephalography (EEG) study employing a visual selective attention task where participants shifted attention to objects characterized by previously rewarded features even if they knew this was counterproductive (Hickey et al., 2010). In line with our results, reward affected vision independent of the pure allocation of endogenous attention but also changed visual saliency directly. Thus, even though reward often leads to attentional allocation to the rewarded stimulus, reward can also act independently of attention (Baldassi & Simoncini, 2011). A recent fMRI study varied attentional demand and reward independently and found that—at least for some subcortical reward-related structures, the ventral tegmental area and ventral striatum—reward anticipation modulated the blood-oxygen-level dependent (BOLD) responses irrespective of attentional load (Rothkirch, Schmack, Deserno, Darmohray, & Sterzer, 2013). This provides one potential mechanism for attention-independent modulation by reward. In sum, there is converging evidence from behavior, animal neurophysiology, and human imaging that rewards can modulate perception akin to visual selective attention without necessarily employing attentional mechanisms.

With the present psychophysical data alone, we cannot distinguish whether the *mechanisms* underlying the attention-like effects of value are identical to attention mechanisms or act through partially distinct circuitry. If attention and reward influence perception similarly and if they indeed act through the same mechanisms, is there any distinction at all? Specifically, is there any conceptual difference between asking observers to attend a percept, to keep it dominant as long as possible (van Ee et al., 2005), or to provide reward for succeeding in doing so? In general, a conceptual distinction can be drawn if attention is understood as *selective attention* in a Jamesian sense, that is as the "withdrawal from some things in order to deal effectively with others" (James, 1890). In contrast to this selectivity, rewards can act through unspecific signals, and reward-based learning then requires relating outcome to stimulus features or to one's own actions. In the present paradigm, and to our knowledge in all rivalry/attention paradigms to date, reward effects cannot easily be distinguished from attention, as the assignment between percept and reward is trivial: There are only two percepts and the mapping between percept and reward is unambiguous. In combination with our findings and with our proposal to objectify the apparently subjective perception in rivalry, this observation may point to a future research direction: To dissociate selective attention from valuation signals, one could—in addition to the perceptual ambiguity in rivalry—add associative uncertainty to the mapping between perceptual dominance and value.

Keywords: attention, reward, binocular rivalry, perception, decision making

Acknowledgments

This work was supported by the German Research Foundation (grants: EI 852/3-1, IRTG-1901, and SFB/TRR-135).

Commercial relationships: none.
Corresponding author: Svenja Marx.
Email: svenja.marx@physik.uni-marburg.de.
Address: Neurophysics, Philipps-University, Marburg, Germany.

References

Alpers, G. W., Ruhleder, M., Walz, N., Mühlberger, A., & Pauli, P. (2005). Binocular rivalry between

emotional and neutral stimuli: A validation using fear conditioning and EEG. *International Journal of Psychophysiology*, *57*(1), 25–32, doi:10.1016/j.ijpsycho.2005.01.008.

Baldassi, S., & Simoncini, C. (2011). Reward sharpens orientation coding independently of attention. *Frontiers in Neuroscience*, *5*, 13, doi:10.3389/fnins.2011.00013.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436, doi:10.1163/156856897X00357.

Brascamp, J. W., & Blake, R. (2012). Inattention abolishes binocular rivalry: Perceptual evidence. *Psychological Science*, *23*(10), 1159–1167, doi:10.1177/0956797612440100.

Breese, B. B. (1899). On inhibition. *Psychological Review*, *3*(1), 1–65, doi:10.1037/h0092990.

Bülthoff, H. H., & Yuille, A. L. (1996). A Bayesian framework for the integration of visual modules. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 49–70). Cambridge, MA: MIT Press.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313, doi:10.1038/nn1194.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink toolbox: Eye tracking with Matlab and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*, 613–617, doi:10.3758/BF03195489.

Deco, G., Rolls, E. T., Albantakis, L., & Romo, R. (2013). Brain mechanisms for perceptual and reward-related decision-making. *Progress in Neurobiology*, *103*, 194–213, doi:10.1016/j.pneurobio.2012.01.010.

Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science*, *20*(6), 778–784, doi:10.1111/j.1467-9280.2009.02360.x.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences, USA*, *105*(5), 1704–1709, doi:10.1073/pnas.0707727105.

Fahle, M. W., Stemmler, T., & Spang, K. M. (2011). How much of the "unconscious" is just pre-threshold? *Frontiers in Human Neuroscience*, *5*, 120, doi:10.3389/fnhum.2011.00120.

Failing, M. F., & Theeuwes, J. (2014). Exogenous visual orienting by reward. *Journal of Vision*, *14*(5):6, 1–9, http://www.journalofvision.org/content/14/5/6, doi:10.1167/14.5.6. [PubMed] [Article]

Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular rivalry: Frontal activity relates to introspection and action but not to perception. *Journal of Neuroscience*, *34*(5), 1738–1747, doi:10.1523/JNEUROSCI.4403-13.2014.

Freeman, W. T. (1994). The generic viewpoint assumption in a framework for visual perception. *Nature*, *368*(6471), 542–545, doi:10.1038/368542a0.

Fries, P., Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences, USA*, *94*(23), 12699–12704.

Glimcher, P. W., & Rustichini, A. (2004). Neuroeconomics: The consilience of brain and decision. *Science*, *306*(5695), 447–452, doi:10.1126/science.1102566.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience*, *30*(33), 11096–11103, doi:10.1523/JNEUROSCI.1026-10.2010.

Hickey, C., & van Zoest, W. (2012). Reward creates oculomotor salience. *Current Biology*, *22*(7), R219–R220, doi:10.1016/j.cub.2012.02.007.

James, W. (1890). *The principles of psychology*. New York: Holt.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304, doi:10.1146/annurev.psych.55.090902.142005.

Lee, D. K., Koch, C., & Braun, J. (1999). Attentional capacity is undifferentiated: Concurrent discrimination of form, color, and motion. *Perception & Psychophysics*, *61*(7), 1241–1255, doi:10.3758/BF03206177.

Lee, J., & Shomstein, S. (2013). The differential effects of reward on space-and object-based attentional allocation. *Journal of Neuroscience*, *33*(26), 10625–10633, doi:10.1523/JNEUROSCI.5575-12.2013.

Lee, S.-H., Blake, R., & Heeger, D. J. (2007). Hierarchy of cortical responses underlying binocular rivalry. *Nature Neuroscience*, *10*(8), 1048–1054, doi:10.1038/nn1939.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254–264, doi:10.1016/S1364-6613(99)01332-7.

Logothetis, N. K., & Schall, J. D. (1990). Binocular motion rivalry in macaque monkeys: Eye domi-

nance and tracking eye movements. *Vision Research*, *30*(10), 1409–1419, doi:10.1016/0042-6989(90)90022-D.

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, *4*(7):2, 539–551, http://www.journalofvision.org/content/4/7/2, doi:10.1167/4.7.2. [PubMed] [Article]

Naber, M., Frässle, S., & Einhäuser, W. (2011). Perceptual rivalry: Reflexes reveal the gradual nature of visual awareness. *PloS One*, *6*(6), e20910, doi:10.1371/journal.pone.0020910.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, *8*(6), 2201–2211.

Ooi, T. L., & He, Z. J. (1999). Binocular rivalry and visual awareness: The role of attention. *Perception*, *28*(5), 551–574, doi:10.1068/p2923.

Paffen, C. L. E., Alais, D., & Verstraten, F. (2006). Attention speeds binocular rivalry. *Psychological Science*, *17*(9), 752–756, doi:10.1111/j.1467-9280.2006.01777.x.

Paffen, C. L. E., & van der Stigchel, S. (2010). Shifting spatial attention makes you flip: Exogenous visual attention triggers perceptual alternations during binocular rivalry. *Attention, Perception, & Psychophysics*, *72*(5), 1237–1243, doi:10.3758/APP.72.5.1237.

Pastukhov, A., & Braun, J. (2007). Perceptual reversals need no prompting by attention. *Journal of Vision*, *7*(10):5, 1–17, http://www.journalofvision.org/content/7/10/5, doi:10.1167/7.10.5. [PubMed] [Article]

Pastukhov, A., Fischer, L., & Braun, J. (2009). Visual attention is a single, integrated resource. *Vision Research*, *49*(10), 1166–1173, doi:10.1016/j.visres.2008.04.011.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. London: Oxford University Press.

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442, doi:10.1163/156856897X00366.

Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*(3), 381–390, doi:10.1016/j.neuron.2006.06.024.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.

Rothkirch, M., Schmack, K., Deserno, L., Darmohray, D., & Sterzer, P. (2013). Attentional modulation of reward processing in the human brain. *Human Brain Mapping*, *35*(7), 3036–3051, doi:10.1002/hbm.22383.

Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *Journal of Neuroscience*, *12*(12), 4595–4610.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61*(5), 700–707, doi:10.1016/j.neuron.2009.01.016.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.

Stănişor, L., van der Togt, C., Pennartz, C. M. A., & Roelfsema, P. R. (2013). A unified selection signal for attention and reward in primary visual cortex. *Proceedings of the National Academy of Sciences, USA*, *110*(22), 9136–9141, doi:10.1073/pnas.1300117110.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, *21*(4), 753–759, doi:10.1016/S0896-6273(00)80592-9.

van Ee, R., van Dam, L. C. J., & Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Research*, *45*(1), 41–55, doi:10.1016/j.visres.2004.07.030.

Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Leipzig, Germany: Voss.

Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*(3–4), 279–292, doi:10.1007/BF00992698.

Zhang, P., Jamison, K., Engel, S., He, B., & He, S. (2011). Binocular rivalry requires visual attention. *Neuron*, *71*(2), 362–369, doi:10.1016/j.neuron.2011.05.035.

Appendix

Dominance durations showed substantial interindividual variability between observers, prompting the use of normalized durations for analysis. However, the direction of the effects was consistent across all observers: For all eight individuals of Experiment 1,

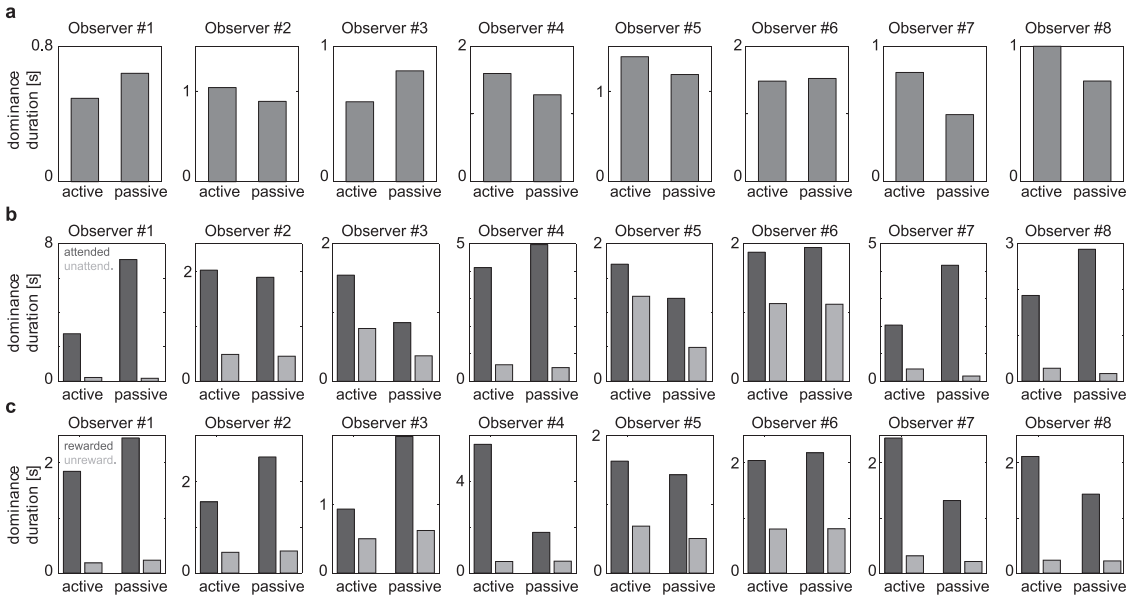

Figure 6. Raw dominance durations for each individual in Experiment 1. (a) No instruction condition, (b) attention condition, and (c) reward condition. Notation as in Figure 2.

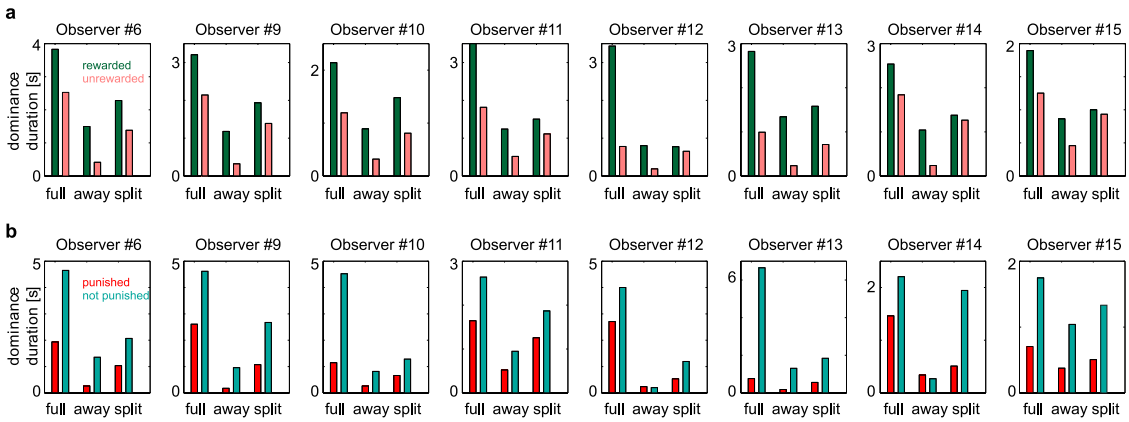

Figure 7. Raw dominance durations for each individual in Experiment 2. (a) Reward part; (b) punishment part. Notation as in Figure 4. Note that Observer 6 of Experiment 1 also participated in Experiment 2.

active and passive conditions yielded similar results (Figure 6), the attended stimulus showed longer dominance than the unattended one (Figure 6b), and the rewarded stimulus dominated longer than the unrewarded one (Figure 6c). Similarly, for Experiment 2, for all levels of attentional load and each individual, the rewarded percept had longer dominance durations than the unrewarded one (Figure 7a). The same held—with the exception of two observers in the away condition where dominance durations were close to floor—for unpunished relative to punished stimuli (Figure 7b). Hence, despite the large interindividual variability in the dominance durations that is typical for rivalry, the effects of attention, reward, and punishment are remarkable robust across observers.

Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *Competitive Visual Processing Across Space and Time*

Competition with and without priority control: linking rivalry to attention through winner-take-all networks with memory

Svenja Marx,[1] Gina Gruenhage,[2] Daniel Walper,[1] Ueli Rutishauser,[3,4] and Wolfgang Einhäuser[1]

[1]Neurophysics, Philipp University of Marburg, Marburg, Germany. [2]Bernstein Center for Computational Neurosciences, Berlin, Germany. [3]Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, California. [4]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California

Address for correspondence: Wolfgang Einhäuser, Philipps-Universität Marburg, AG Neurophysik, Karl-von-Frisch-Str. 8a, 35032 Marburg, Germany. wet@physik.uni-marburg.de

Competition is ubiquitous in perception. For example, items in the visual field compete for processing resources, and attention controls their priority (*biased competition*). The inevitable ambiguity in the interpretation of sensory signals yields another form of competition: distinct perceptual interpretations compete for access to awareness. Rivalry, where two equally likely percepts compete for dominance, explicates the latter form of competition. Building upon the similarity between attention and rivalry, we propose to model rivalry by a generic competitive circuit that is widely used in the attention literature—a winner-take-all (WTA) network. Specifically, we show that a network of two coupled WTA circuits replicates three common hallmarks of rivalry: the distribution of dominance durations, their dependence on input strength ("Levelt's propositions"), and the effects of stimulus removal (*blanking*). This model introduces a form of memory by forming discrete states and explains experimental data better than competitive models of rivalry without memory. This result supports the crucial role of memory in rivalry specifically and in competitive processes in general. Our approach unifies the seemingly distinct phenomena of rivalry, memory, and attention in a single model with competition as the common underlying principle.

Keywords: psychophysics; modeling; vision; binocular rivalry; attention; winner-take-all network

Introduction

When confronted with complex and potentially ambiguous input, human sensory systems have to deal with two forms of competition. First, different items in the visual field compete for processing resources; second, different possible interpretations of the sensory signal compete for perceptual awareness.

Attention as biased competition

The first form of competition is typically resolved by attention, enhancing one stimulus at the expense of the other.[1] This is most evident in the framework of *biased competition*,[2] where attention corresponds to resolving competition by setting biases (i.e., controlling priority) according to task demands.[3] Biased competition has become one of the most influential attention models[4,5] and is supported by ample physiological evidence: when two stimuli are brought into a cell's receptive field (RF), of which one alone would drive the cell and the other would not, the cell's response to the combined stimulus falls in-between the two individual responses, as a consequence of competition. When, however, one stimulus is attended, the neuron's response quickly behaves as if only the attended stimulus would be present in the RF; that is, competition is biased in favor of the attended stimulus.[6]

Attention and memory

Attention and visual working memory are tightly linked.[7,8] For example, items held in working memory can interfere with attentional selection and vice versa.[9–12] Consistent with such evidence, an

doi: 10.1111/nyas.12575

early formalization of the biased competition idea, Bundesen's theory of visual attention[13] (TVA) and its later neural implementation (neural theory of visual attention, NTVA),[14] describes attention as a race of competing items for visual short-term memory. TVA formalizes the interplay of Broadbent's two mechanisms of attention:[15] *filtering*, the mechanism for the selection of items, and *pigeonholing*, the mechanism to allocate evidence to categories. Since filtering represents the probability of an *item* to be selected, while pigeonholing represents the probability of a *category* to be selected, their complementary functions parallel the aforementioned two forms of competition: filtering resolves competition between items; pigeonholing resolves competition between different categories, including different perceptual interpretations. NTVA[14] provides a neuronal implementation of these mechanisms that is consistent with physiological data. In NTVA, filtering and pigeonholing are related to specific neural mechanisms, namely the allocation of RFs to select elements and of gain control to select categories. In an extension of the NTVA, a Poisson counter model is used to explain how during visual identification mutually confusable stimuli can be resolved.[16] It implies that while the stimulus is analyzed, temporary categorizations are made at a constant Poisson rate. The response is then based on the category that was chosen most frequently. Thereby, the Poisson counter model provides a mechanism by which the interplay of attention and memory can resolve competition between distinct perceptual interpretations of a visual stimulus.

Rivalry as a model for competition

The second form of competition, the competition of perceptual interpretations for awareness, is unavoidable during natural vision. Because of the inherent ambiguity when mapping the outside world on the receptive surface,[17] prior knowledge is needed to infer the most likely interpretation. Such prior knowledge can manifest itself in terms of fixed rules about object structure—with Gestalt laws as a prime example[18]—or formalized in terms of Bayesian prior distributions,[19–23] which may be flexibly adapted to environmental and motor constraints.[24] On the basis of sensory input alone, many perceptual alternatives may be equally likely, but the combination of this likelihood with the prior assumptions allow the sensory system to arrive at a unique interpretation of the world. If no sufficiently strong prior information is available to resolve the ambiguity in the input, the system will nonetheless perceive one unique interpretation at any point in time, but the dominant interpretation alternates over time. This phenomenon is referred to as rivalry, which can be induced either through bi- or multistable figures, such as geometrical figures that alternate in three-dimensional interpretation,[25,26] figure-ground reversals,[27,28] or overlaid patterns that alternate between compound and constituents,[29,30] or as "binocular rivalry," when two sufficiently distinct patterns are presented to either eye[31] (for review, see Ref. 32). Most forms of rivalry have several properties in common.[33,34] The times that a certain precept dominates are distributed with a leptokurtic (heavy-tailed) distribution[35] and respond in a well-defined manner to changes in input strength (Levelt's propositions[36]).

The role of memory in rivalry

If the stimulus is removed ("blanked") for a considerable duration (>500 ms) during a rivalry task, the probability that the same perceptual interpretation reemerges after the blank increases substantially.[37,38] Thus, blanking stabilizes the percept. In contrast, for short blank durations (<500 ms), the percept tends to destabilize and thus the alternative percept is more likely to emerge after the blank than expected by chance.[37] The time course of the blanking effect is reminiscent of a recently proposed "third stage" in visual working memory encoding that protects an item from deletion when its processing takes longer than the completion of a competition epoch,[39] and it is tempting to speculate that the stabilization of the blanked percept is a consequence of such protective maintenance.

Stabilization of the percept across extended periods of blanking indicates that a form of memory—in this case, the dominant percept before onset of the blank period—plays a role in rivalry. Additional evidence for the role of memory in rivalry comes from experiments with tri-stable rivalry (i.e., a stimulus with three possible percepts). In these experiments, the sequence of states is not Markovian (i.e., previous percepts influence processing of the current perceptual state[40]). For brief intermittent presentations, the dominant percept is location specific: at a given location of the visual field, the same percept is

Ann. N.Y. Acad. Sci. xxxx (2015) 1–16 © 2015 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals Inc. on behalf of The New York Academy of Sciences.

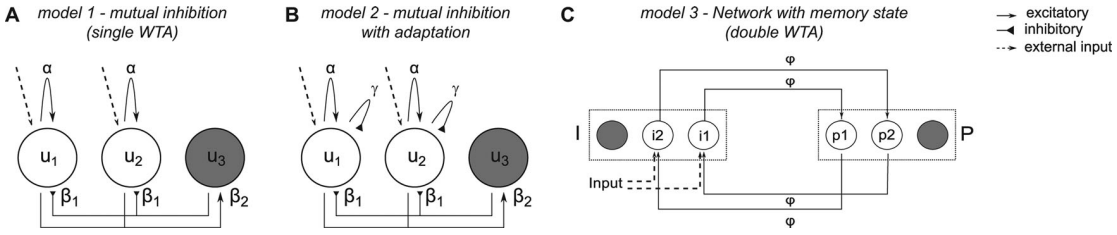

Figure 1. Network models: the three models tested in this study. (A) Model 1: a single WTA circuit; each excitatory unit is recurrently coupled to itself with weight α and to the inhibitory unit with weight $\beta 2$. In turn, the inhibitory unit is coupled to both excitatory units with weight $\beta 1$, but not to itself. Input is applied to both excitatory units, and the perceptual states are recorded directly from these units. (B) Model 2: identical to model 1, except that both excitatory units are adapting (see Methods in the Supporting Information for details). (C) Model 3: two WTA circuits, as used in model 1, are coupled by connecting their excitatory units across circuits; all connections between the circuits have the same weight ϕ, but feedback connections cross between the two sets of neurons representing different states. Input is applied to map I, and percepts are recorded from map P.

dominant at onset after blanking throughout.[41] Although these biases in onset rivalry are highly variable between observers, they remain stable within the same individual over weeks. This suggests involvement of long-term memory. Taken together, blanking, the non-Markovian property of tri-stable rivalry, and the observer-specific location bias of onset rivalry show that rivalry is influenced by a number of memory processes that operate on a variety of time scales.

A common framework for rivalry and attention as competitive processes—winner-take-all circuits

In neuronal circuits, competition is frequently implemented by winner-take-all (WTA) circuits. WTA behavior emerges if a population of excitatory neurons is recurrently connected to itself and shares a common inhibitory signal[42,43] with sufficiently high gain. Such recurrent connectivity is a building block of neocortical circuitry[44–46] and is readily implemented in neuromorphic hardware.[47] WTA networks can model arbitrary state machines,[48] states can remain in the absence of input, and state transitions can be triggered by external input given the current state.

WTA circuits have frequently been used in models of attention. The output stage of the saliency map,[49] which must select a winning location, is typically implemented as a WTA circuit. More deeply, attention models can be built by cascading WTA circuits[50] or by implementing WTA mechanisms between visual filters.[51] In a related architecture, Hahnloser *et al.*[52] argue that a recurrently coupled map alone cannot implement attention to a region of the map, but rather propose an excitatory reciprocal coupling between the map and a "pointer" map, whose neurons are more broadly tuned in space.

Here, we propose to exploit the structural similarity between rivalry and attention as forms of competition and present a WTA model of rivalry. We start with a generic WTA model rooted in neocortical physiology[43] and test the extent to which it replicates the dominance distributions and Levelt's propositions as main hallmarks of rivalry. We then demonstrate that the required memory state emerges from the network's dynamics. The model predicts interactions between blanking duration and input strength that we subsequently test experimentally.

Materials and methods

Modeling

Our aim is to construct a comprehensive model of rivalry that replicates the three key features common to all rivalry processes: leptokurtic dominance distributions, Levelt's propositions, and the role of memory, in particular for the phenomena related to stimulus blanking. We propose that a network consisting of two coupled WTA circuits exhibits all these features. For comparison, we also analyze representatives of other modeling approaches that have been proposed for rivalry by embedding them in the WTA framework.

Specifically, we here compare three models of rivalry (Fig. 1), which serve as prototypes for broad classes of rivalry models: first, networks of self-exciting units with mutual inhibition (model 1); second, the same network augmented with an adaptation mechanism in the excitatory units (model 2); third, our new approach, two coupled networks that

implicitly form a memory state (model 3). In this section, the models are outlined; for a detailed mathematical description, implementation details, and parameter choices (Table S1), the reader is referred to the methods described in the Supporting Information.

The fundamental circuit for all three cases is a single WTA network. This network consists of two excitatory units and one inhibitory unit. Each excitatory unit excites itself and projects to a global inhibitory unit, which, in turn, projects back to both excitatory units. Units are mean-rate approximations of the activity of a group of individual neurons, and activity is modeled with respect to average rates rather than individual spike times. Since there is no explicit mapping from the time units in the simulation to real time, we consistently use a unit of 1000 steps of the Euler integration (see Methods in the Supporting Information) as the time unit when reporting modeling data and parameters. Simulations were performed in Matlab (The MathWorks, Natick, MA, USA) on the basis of the code available at http://www.ini.uzh.ch/~urut/DFAWTA[48]; each condition (combination of input currents) was simulated five times with different random noise patterns.

Model 1: mutual inhibition. For model 1, the fundamental WTA circuit is considered in isolation (Fig. 1A). Even though the inhibitory unit is modeled explicitly, this network corresponds to a network with self-exciting units that mutually inhibit each other,[53] except for the delayed inhibition caused by the inhibitory unit. To probe the model, input currents with Gaussian noise are applied to both excitatory units. The value of the input current represents the sensory input corresponding to one of the possible percepts. The activity level of the excitatory units determines which percept is currently dominant. The perception belonging to one unit is considered dominant whenever the activity of the respective unit exceeds double the activity of the other unit. The remainder constitute transition periods, which were not considered further in the present context, neither experimentally nor in modeling.

Model 2: mutual inhibition with adaptation. Many models of rivalry assume a form of "fatigue" or habituation; that is, if a percept has been dominant for some time, its representation fatigues and thus the other percept becomes dominant. On a neuronal level, the equivalent of such fatigue is neuronal adaptation. In model 2, we implement adaptation by adding an additional term to each excitatory unit (Fig. 1B). Otherwise, model 2 is identical to model 1. This results in a model of mutual inhibition with adaption, akin to the model used in Ref. 54.

Model 3: two coupled circuits, implicit memory state. When the external input is removed from a single WTA network, its activity relaxes back to zero and it therefore has no memory. As this is in conflict with experimental evidence, in particular with the increased survival probability of a percept after prolonged stimulus removal (*blanking*), we consider a third model that implicitly implements a memory state in its dynamics. To do so, we couple two of the WTA circuits as used in model 1 (Fig. 1C). One of the circuits (I, with excitatory units i1 and i2) represents the input layer, while the other (P, with excitatory units p1 and p2) represents the perception layer, from which activity is "recorded." Importantly, the feedforward connection from the units of the input circuit (i1 and i2) project to the corresponding perception units (p1 and p2, respectively), while the feedback projections map onto the input corresponding to the alternate percept (p2 to i1 and p1 to i2). This network can maintain its current winner (state), even if the external input is removed.[43] Such persistent activity that is maintained in the absence of external input endows the network with a memory, because the state active during the removal of external input is maintained (see Ref. 48 for details). In the context of rivalry, this makes the percept after the blank period has ended conditional on the state (i.e., the percept) before stimulus removal.

Model input; simulation of blanking. In typical simulation runs, constant input with Gaussian noise is supplied to both input neurons (model 1/2: u1 and u2; model 3: i1 and i2) of the network. Stimulus strength is set by the mean current applied. For simulating stimulus removal and reappearance (blanking), we in addition model a sensory neuron that is located upstream (i.e., lower in the visual hierarchy) to the rivalry circuit. This is done by modulating the injected current accordingly: at stimulus onset, the current transiently rises to thrice the sustained value, followed by a rapid exponential decay to the sustained value (α function $t/\tau_\alpha \times \exp(1 - t/\tau_\alpha)$,

Ann. N.Y. Acad. Sci. xxxx (2015) 1–16 © 2015 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals Inc. on behalf of The New York Academy of Sciences.

with time constant $\tau_\alpha = 0.025$). Stimulus offset is modeled by the current decaying to baseline level (activity of 0.1) in the shape of a hyperbolic tangent (half life: 0.080). Stimulus durations were fixed to 1.0 and blank duration was varied between 0.1 and 2.0 in steps of 0.1 (all times in units of simulated time as defined above).

Analysis of modeling data. To mimic an instruction in which observers report exclusive dominance, we define a percept to be dominant, whenever the respective unit's activity exceeds the other unit's activity by at least a factor of 2. Periods in which none of the percepts are dominant according to this definition are defined as transition periods. We define the dominance duration of a percept as the time from its onset to its offset irrespective of whether the same or another percept follows. For further analysis, we excluded every first and last dominance duration since these are not restricted by the network dynamics but by onset and offset of the simulation. Since dominance durations within a trial cannot be expected to follow a Gaussian distribution, we quantify the distribution of dominance durations in each condition by the distribution's median. For comparison to the experimental data and among models, we normalize all dominance durations by the condition with highest input to both eyes. For quantification of the dominance of one state, which is independent of the respective median dominance duration, all dominance durations of this percept obtained in one simulation period were added. Denoting the resulting sums for the two percepts as D_1 and D_2, respectively, we define the relative dominance as:

$$D_{\text{rel}} = \frac{D_2 - D_1}{D_2 + D_1}. \quad (1)$$

Analysis of blanking. To quantify the effect of blanking, we define a percept's survival probability as the number of blanking intervals for which the dominant percept before and after the blank was identical divided by the total number of blanks. To closely match our experimental instructions and accounting for human reaction time, the dominant percept for each presentation interval is measured 0.1 time units after the onset of the respective presentation. Similarly, we define a switch probability across blanks as the number of blanking intervals for which the dominant percept before and after the blank was different. Since, occasionally, observers (and models) do not report a dominant percept during a presentation interval, survival probability and switch probability do not necessarily add up to 1 and the difference of their sum and 1 quantifies such failures to report.

Unlike in behavior, the recorded units signal a percept even during the blanking period. Such "hallucinations" can easily be suppressed by an additional downstream gating mechanism that allows a percept only to get to awareness, if any input is present. Here, we do not model this explicitly, but merely ignore the period of the blank as such for further analysis.

Behavioral experiments

Observers. Five observers (age: 21–26 years; four female) participated in experiment 1; five (age: 24–26 years, three female) participated in experiment 2. Two observers participated in both experiments. All had normal or corrected-to-normal vision and were naive to the purpose of this study. All gave written consent before the experiment. The experiments conformed with the Declaration of Helsinki and were approved by the local ethics committee (Ethikkommission FB04).

Stimuli. Each eye was presented one sinusoidal grating (3.4 cycles per degree; mean luminance = 25.1 cd/m^2), oriented $+45°$ in one eye and $-45°$ in the other. Gratings had full contrast in a circular patch of 0.3° radius outside of which contrast fell off with a Gaussian profile (SD = 0.11°). To facilitate fusion, the patch was surrounded by an alignment annulus (radius = 1°, width = 0.06°) of white noise of the same mean luminance.

The contrasts of the gratings were adjusted to each individual's detection threshold, which was defined as the 75% correct level as identified by a 2AFC QUEST[55] procedure. For none of our observers was there any significant difference in threshold between their eyes. Similar to the methods described in Ref. 56, the lowest contrast used was 0.75 $\log_{10}$ above this threshold; the highest contrast was 100% Michelson contrast, and the four contrast levels in between were logarithmically spaced. This defines six contrast levels in each individual, which hereafter are referred to as contrast 1 to contrast 6.

Apparatus. Stimuli were presented separately to each eye through a stereoscope on two 21-in.

Samsung Syncmaster CRT screens at a viewing distance of 30 cm by an Optiplex Dell computer running Matlab with a Psychophysics toolbox extension (http://psychtoolbox.org[57,58]). Each screen had a resolution of 1280 × 1024 pixels, a refresh rate of 85 Hz, and was γ corrected to achieve the same linear mapping of pixel values to stimulus luminance in both eyes. Eye position was monitored throughout the experiment by an Eyelink-1000 (SR Research, Osgoode, ON, Canada) eye-tracking device, and the device was calibrated at the onset of each trial; although in the context of this study, eye-tracking data are not considered further.

Procedure: experiment 1. After identifying the individual's detection threshold, observers performed 72 2-min experimental trials, chunked into 12 blocks of 6 trials. For these trials, each of the six contrast levels was combined with any other level and this was done for both possible assignments of grating orientation to eye ($72 = 6 \times 6 \times 2$). Order of these trials was random. To control for changes in overall behavior and for normalization purposes, between experimental trials 3 and 4 of each block, an additional control trial that used full-contrast stimuli in each eye (contrast level 6) was inserted. For these 12 control trials, we did not observe any significant change for any individual in any measure of interest. To have an equal amount of experimental trials in each condition, in the main analysis control trials were used only for normalization of dominance durations and not analyzed otherwise.

Participants reported their current percept by pressing and holding one of two buttons on a game pad. They were instructed to fixate the gratings throughout an experimental session and only report a percept when it appeared clearly dominant and refrain from any button press, when both percepts appeared about equal.

Procedure: experiment 2. In the second experiment, each trial started with 90 s of continuous presentation of the rivalrous stimuli, followed by a 180-s period of intermittent presentation and another 90 s of continuous presentation. In the intermittent part, stimuli were repeatedly presented for 0.5 s and removed for a fixed period of blanking. Across trials, three different contrasts (levels 2, 4, 6 to both eyes) and four different blanking durations (0.5, 1, 2, and 4 s) were used, resulting in 12 trials per participant. Since presentation duration was short, participants were instructed to press the button indicating their percept only once during each presentation period or shortly afterward and press no button during the blanking periods. Otherwise, the procedure was identical to experiment 1.

Data analysis. Akin to the analysis of the modeling data, we define each period during a trial in which exactly one button was pressed as dominance period for the respective percept, and other periods as transition periods. In each trial, the time before the first button press and the last dominance period, which the trial end interrupted, were excluded from analysis. To normalize for interindividual differences in group analyses and comparisons to modeling, all dominance durations were divided by the median dominance duration of the 12 control trials with full contrast to both eyes. Definitions of relative dominance, dominance durations, and switch rate are then analogously defined to analyzing modeling data.

Since we did not observe any differences between grating orientation, we pooled dominance durations across orientations. For the analysis of relative dominance and switch rate, we separate by left and right eye, resulting in an effective 6 × 6 design with six levels for left-eye contrast and six levels for right eye contrast. Since we did not observe any eye to be preferred for any observer, analysis of median dominance durations, where the contrast to the eye whose dominance duration is considered (ipsilateral eye) has to be distinguished from the other eye (contralateral eye), is pooled across eyes.

Results

To compare different computational models of rivalry, we simulated three different networks of increasing complexity. Model 1 is a single WTA circuit with mutual inhibition and noisy input; model 2 adds adaption; and model 3, by combining two WTA circuits, an implicit memory state. We assessed each model according to three hallmarks of rivalry: dominance distributions, Levelt's propositions, and the effects of periodical stimulus removal, and compared the predictions to new experimental data.

To simulate rivalry, noisy external inputs were provided to units u1 and u2 in the single WTA cases or to units i1 and i2 of the circuit I in the double WTA case, respectively. Input strength was modeled

Ann. N.Y. Acad. Sci. xxxx (2015) 1–16 © 2015 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals Inc. on behalf of The New York Academy of Sciences.

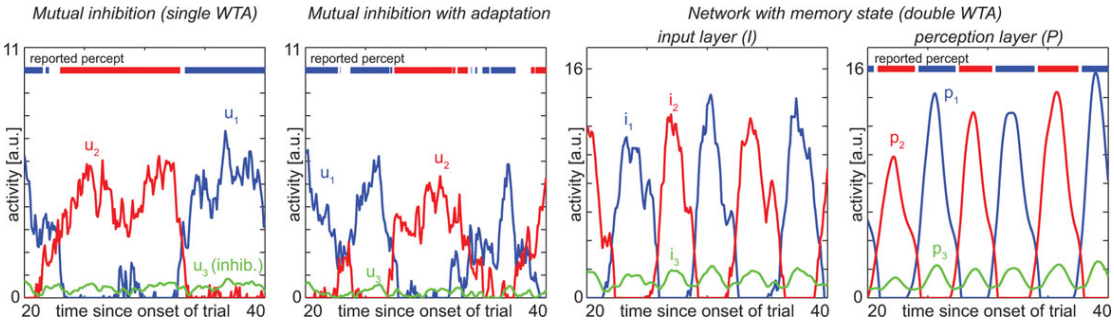

Figure 2. Raw activity. Activity traces for the three models when noisy input (strength 6.5 for models 1 and 2 and strength 5.5 for model 3) is applied to both eyes (see Methods in the Supporting Information for units). All models show bistable behavior, with the excitatory units (blue, red) alternating in dominance. The currently dominant percept, according to the definition used throughout, is indicated by the red and blue bars on top of each plot for models 1 and 2, and for the percept units of model 3. Green trace represents the activity of the inhibitory unit.

by adjusting the mean of the input currents and adding Gaussian noise of constant standard deviation. In the behavioral experiments, input strength was given by the log contrast of the stimulus relative to the individual's threshold.

Example data, dominance durations, and dominance

In the experimental data, all observers experienced rivalry, and valid dominance data (exactly one button pressed, no interruption by trial end) were obtained for 85.8% of the total time (range across observers and conditions: 76.3–95.4%). The remaining time consists of periods of mixed percepts, transition periods, and discarded data at the beginning and end of the trial.

All three model networks show bistable behavior (Fig. 2), which allows us to define periods of perceptual dominance. Akin to the experimental instruction to report a percept only if it is clearly dominant, we define a percept to be dominant in simulation if its unit's activity exceeds twice the activity of the other unit. Using this criterion and the same end-of-trial exclusion as in the experimental data, we can—again averaged over all conditions—identify a dominant percept for 68.6% (range: 0–95.6%) of time for model 1, 87.3% (81.8–91.1%) of time for model 2, and 89.9% (88.8–91.1%) for model 3. Except model 1, in which for some asymmetric input conditions one percept is dominant throughout, yielding no valid data, the amount of data usable therefore is comparable to the experimental situation.

Distribution of dominance durations

To address the most typical rivalry situation in which both percepts are about equally strong, we first consider the conditions in which stimuli of the same strength were presented (symmetric input). This was done by injecting currents of the same mean into each input unit (simulation) or presenting stimuli of the same contrast to each eye (experiment). While absolute values of dominance durations and the spread of distributions typically vary largely between individuals and rivalry type,[59] nearly all rivalry types exhibit leptokurtic (heavy-tailed) dominance durations. Our experiment 1 confirms this tendency, with showing leptokurtic dominance distributions for all contrast levels tested (minimal kurtosis: 6.3, with values larger 3 implying leptokurtic distributions; Fig. 3). Across all symmetric input conditions, all models show a kurtosis of larger 3, with minimum values over conditions of 5.1 (model 1), 9.9 (model 2), and 4.2 (model 3). However, model 1 and 2 show an abundance of short dominance durations (Fig. 3, left panels and respective insets) as compared to model 3 (and to a lesser extent our experimental data). Nonetheless, all models qualitatively replicate the leptokurtic distribution of dominance durations that is common to nearly all rivalry phenomena.

Levelt's propositions

For a more detailed analysis of the dependence of rivalry on input strength, we consider situations in which both input strengths are varied independently. For a broad range of rivalry phenomena, the dependence of dominance, dominance durations,

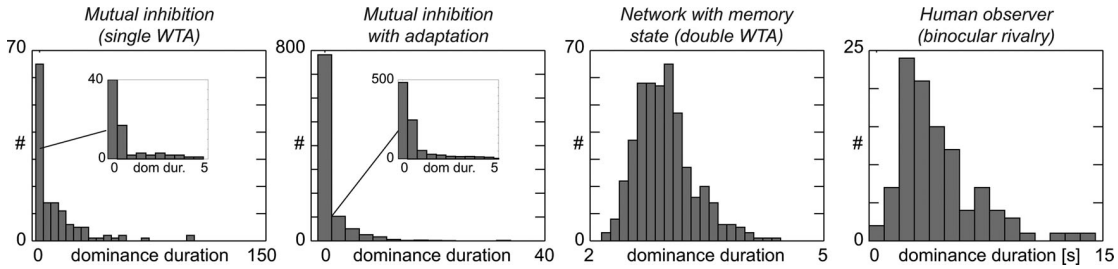

Figure 3. **Dominance distributions. Example distributions of dominance durations for the three models and experiment 1 for a medium input strength. Modeling data are based on a single simulation run and experimental data on a single individual. Dominance durations are pooled over both percepts. Insets depict finer resolution for the left-most bin (model 1) or two left-most bins (model 2), corresponding to five time units to ease comparison with model 3.**

and switch rates on the two input strengths then follows certain rules, typically referred to as Levelt's propositions.[36] Here, we test the extent to which our models reproduce Levelt's propositions and again compare the data to our experimental observation.

Levelt's first proposition: increase of stimulus strength in one eye will increase the predominance of the stimulus. For all models and the experimental data, we calculate the relative dominance of each combination of input strengths (input currents or contrast levels, respectively). By definition, a relative dominance of 0 corresponds to equal dominance of either percept, positive values dominance of percept 2 or right eye, negative values of percept 1 or left eye. Consistent with Levelt's first proposition, we find relative dominance to increase when input to the right eye or the corresponding input unit u2 or i2 is increased, to decrease when input to the left eye (or unit u1 or i1) is increased, and to fall around 0 when the input to both is the same (Fig. 4A). Quantitatively, however, there are substantial differences: model 1, the single WTA circuit, only has a narrow band around equal input strength in which dominance does not get stuck at the extreme. When input is applied asymmetrically, there is no mechanism to release the nondominant state from suppression as soon as noise becomes negligible. Adaption in model 2 counters this effect, and the extremes are approached in a more shallow fashion. Importantly, a qualitatively very similar behavior is observed for the double WTA network of model 3, even though there is no explicit adaptation mechanism at the level of an individual unit. The experimental data also show a broad range and smooth variation as do models 2 and 3. Unlike those models, however, experimental data reach the extremes of full dominance, while these models do not exceed a relative dominance of about ± 0.5 (i.e., one input dominating for 75% of time) for the input range tested. Nonetheless, the double WTA (model 3) and the single WTA with adaptation (model 2) similarly capture the smooth transition of relative dominance from one eye to the other when input strength is changed.

Levelt's second proposition: increase of stimulus strength in one eye will not affect the mean dominance time for the same eye. Of the four Levelt's propositions, the second is arguably the most counterintuitive and has been challenged recently.[56,60] The resulting revised version of this proposition states that "changes in contrast of one eye affect the mean dominance duration of the highest contrast eye."[56]

To analyze our models and data with respect to Levelt's second proposition, we plot the dominance duration of one eye/input unit as a function of the input of this (ipsilateral) eye/input unit and the other (contralateral) eye/input unit (Fig. 4B). For analysis, we fix one input strength and vary the other (i.e., we proceed either along rows or columns of the panels in Fig. 4B). We then can consider either the dominance durations of the "fixed" input or of the "variable" input. For illustration, Figure S1 shows some of the data of Figure 4(B) in this representation.

The modified version of Levelt's second propositions predicts that the median dominance duration of the percept receiving higher input strength should vary most. In the extreme cases of highest

Ann. N.Y. Acad. Sci. xxxx (2015) 1–16 © 2015 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals Inc. on behalf of The New York Academy of Sciences.

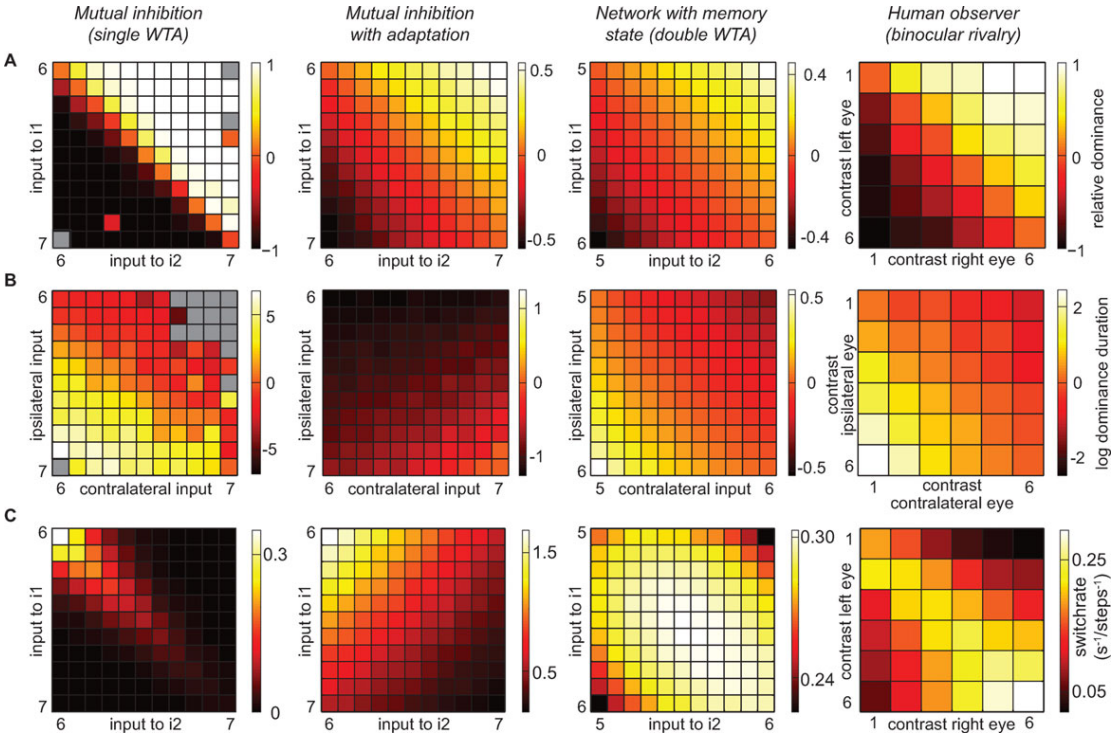

Figure 4. Levelt's propositions. (A) Levelt's first proposition tested for the three models and data of experiment 1; relative dominance is color coded individually per panel. In the panel for model 1, some simulations are stuck within the same state throughout, and—as for all analysis the last period is excluded—no data is available, indicated in gray. (B) Levelt's second proposition: log dominance duration for one eye (ipsilateral eye) while input strength to this eye and to the other eye (contralateral eye) are varied independently. Data are collapsed over both eyes (left/right) or units (i1/i2, p1/p2). Log scale is used for illustration, and correlations are computed on the original data. (C) Levelt's third and fourth propositions: dependence of switch rate on input strength to either eye.

and lowest fixed input strength, this would result in decreasing dominance durations of the percept receiving fixed input strength in the first case and in increasing dominance durations of the percept receiving variable input strength in the latter, while the dominance durations of the other percepts remain stable.

At lowest fixed input strength, all networks show qualitatively the same behavior as the experimental data (Fig. S1A); namely, the median dominance duration of the percept receiving variable input increases with increasing input strength, while the median dominance durations of the percept receiving fixed input strength stays largely constant. The simulated and experimental data are thus in line with the modified version of Levelt's second proposition. However, when the fixed input strength is increased, the single WTAs behaves differently from the double WTA model. Only the double WTA model is consistent with the experimental data (Fig. S1B and C).

To quantify this, we compute correlations between fixed input strength and median dominance durations for each level of variable input strength and vice versa (i.e., we compute correlations within either each row or column of the panels in Fig. 4B). For experiment and model 3, correlations between input strength and median dominance durations of the percept receiving variable input are strictly positive and significant for all input strengths (double WTA: all $r(9) > 0.96$, all $P < 3.7 \times 10^{-6}$; experiment: all $r(4) > 0.90$, all $P < 0.019$)). In contrast, dominance durations of the input receiving fixed input strength and input strength to the variable input are correlated negatively for all input strengths (double WTA: $r(9) < -0.96$, $P < 1.82 \times 10^{-5}$)); experiment: $r(4) < -0.89$, $P < 0.014$)). The single WTA without adaptation (model 1) still trends to a negative correlation for the fixed input and the positive correlation for the variable input, even though not always significant (fixed: $r(9) < -0.50$,

$P < 0.11$; variable: $r(9) > 0.61$, $P < 0.045$). In contrast, the single WTA with adaptation (model 2) shows positive correlations between median dominance durations and input strength for both, the percept receiving variable and fixed input strengths (all $r(9) > 0.85$, all $P < 0.00089$), which is—for the variable input—the exact opposite of the experimental observation. Hence, models 1 and 3 replicate the modified version of Levelt's second proposition, whereas a single WTA with adaptation (model 2) shows qualitatively different behavior.

Levelt's third proposition: increase of stimulus strength to one eye will increase the alternation frequency. Levelt's third and fourth proposition are closely linked. Both make predictions on the alternation frequency (here: switch rate), when input strength is varied. The switch rate is the number of switches in dominance per unit time (simulation time or seconds). Every transition from one percept to the other or from one percept to the same percept when there was a transition time in between (same-state transition) is considered a switch for the present purpose.

Levelt's third proposition states that when input strength or contrast to one unit/eye was fixed, increasing input strength/contrast to the other would result in a higher switch rate. In its revised version,[60] the proposition instead states that switch rate is "maximal at and symmetric around equi-dominance." Our experimental data, which reach up to high-contrast levels, confirm the revised version of the proposition (Fig. 4C, right). Model 1 replicates this property, but switch rates rapidly drop to 0 when leaving equi-dominance (Fig. 4C, left). The single WTA with adaptation (model 2), in turn, does not show the symmetry around equi-dominance (Fig. 4C, second panel). In contrast, the double WTA (model 3) shows a distribution of switch rates that is symmetric (Fig. 4C, third panel) and maximal around equi-dominance. Hence, only model 3 qualitatively captures the revised version of the third proposition and is in line with the experimental data.

Levelt's fourth proposition: increase of stimulus strengths in both eyes will increase the alternation frequency. This proposition predicts an increase of switch rate when stimulus strength is increased simultaneously in both units/eyes, which would be reflected by an increase of switch rate along the diagonals in Figure 4(C). Again, models 1 and 2 deviate qualitatively from this prediction by showing a decrease along the diagonal toward increasing input strength (model 1: $r(9) = -0.94$, $P = 2.2 \times 10^{-5}$; model 2: $r(9) = -0.98$, $P = 1.0 \times 10^{-7}$). In contrast, model 3 qualitatively captures the increase with increasing input strength, which we also observe in our experimental data (model 3: $r(9) = 0.67$, $P = 0.024$; experimental data: $r(4) = 0.97$, $P = 9.4 \times 10^{-4}$).

In sum, even though models 1 and 2 capture some aspects of rivalry, only model 3 is—at least on a qualitative level—in line with the experimental observation.

Blanking

Another key phenomenon of rivalry is blanking: after the stimulus is removed intermittently for a sufficiently long time (>500 ms), the percept stabilizes,[37,38] whereas it destabilizes when the blanking duration is shorter than about 500 ms.[37] Stability here means that the same percept is dominant before and after the blank (see Methods). Blanking is an example of the involvement of memory in rivalry. The double WTA model (model 3) has memory for the current percept (its state). We next test this model's ability to replicate the main features of blanking. Experimentally (experiment 2) and in simulation (Fig. 5A), we vary blanking duration and input strength (Fig. 5B). Example traces of activity in model 3 already indicate that the model may replicate the tendency for longer blank durations to lead to more stabilization (Fig. 5C) in line with the experimental example (Fig. 5D).

Quantitatively, we investigate blanking with respect to survival probability, the number of times a percept remerges after the blank is divided by the number of all blanks (Fig. 6A). This number would be 0 if percepts perfectly alternated, 1 if there was the same percept always present, and 0.5 if alternations were random (as there is no bias to either percept in simulation nor experiment). As a consequence of the definition of dominance in simulation and the instruction to only report a percept when it was clearly dominant, a dominant percept is not always identifiable (especially for models 1 and 2) during the presentation. Hence, we also analyze switch probability as the fraction of blanks after which the other precept reemerges after a blank (Fig. 6B). The difference between 1 and the sum of

Ann. N.Y. Acad. Sci. xxxx (2015) 1–16 © 2015 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals Inc. on behalf of The New York Academy of Sciences.

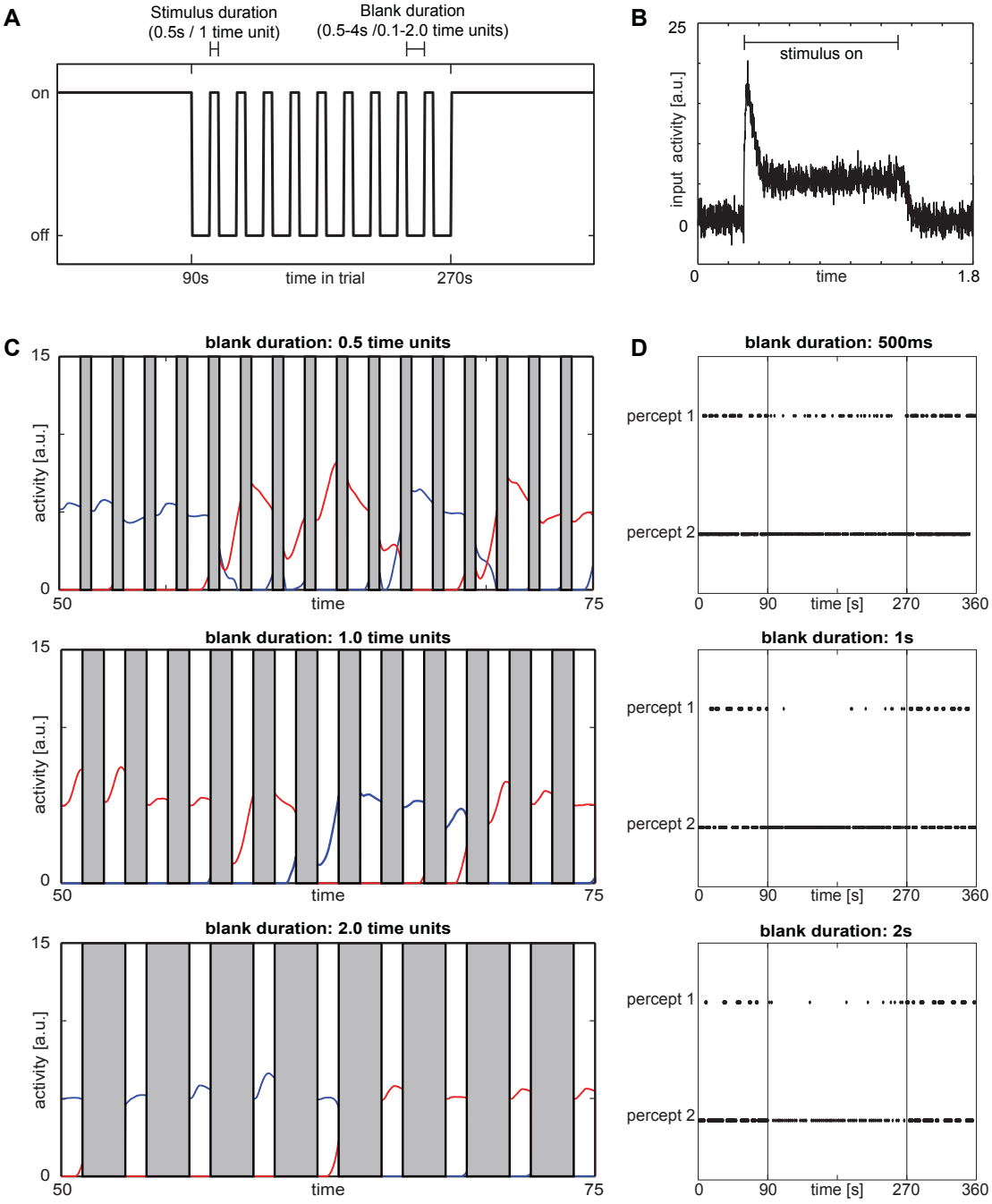

Figure 5. Blanking, model output. (A) Time course of an experimental blanking trial. Blank intervals are not to scale. (B) Input function for modeling blanking in the models; typical example with added noise. At onset there is a steep rise with exponential decay to the sustained activity; at offset a smooth relaxation to baseline. (C) Example traces of model 3 for three different blanking durations. Blue and red curves correspond to neurons p1 and p2, respectively. (D) Experimental data for blanking in a single subject. Percept changes more frequently in the absence of blanks ($<$90 s, $>$270 s), and stabilization depends on blank duration.

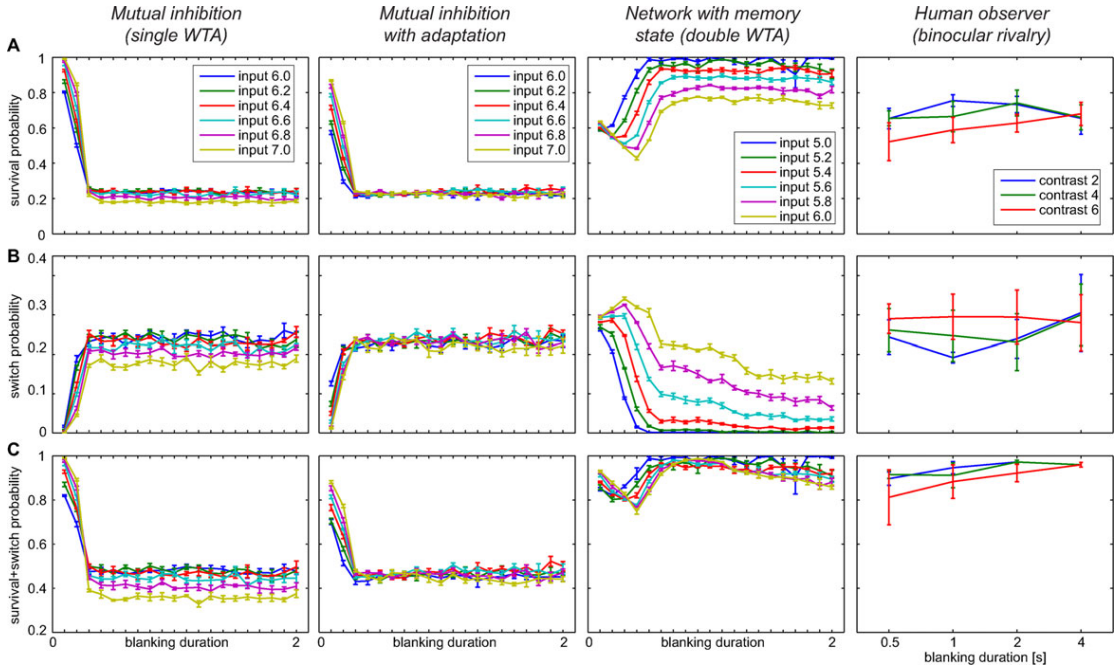

Figure 6. Blanking, model results. (A) Survival probability, (B) switch probability, and (C) their sum for the three models and the data of experiment 2. Different line colors indicate different input strengths (consistent within each column as given in the top-row panels).

switch and survival probability (Fig. 6C) provides the fraction of unidentifiable transitions through a blanking period. Not surprisingly, the two single WTA models (models 1 and 2) do not replicate the blanking phenomenon. Once the input decayed (cf. Fig. 5B), no information about the preceding state is left, and switch and survival probability are similar (Fig. 6A and B, left columns). In addition, there are many situations (up to 67.2%; Fig. 6C, left columns) in which the presentation time does not allow for a clear dominant percept to emerge after a blink. In contrast, the double WTA model (model 3) replicates the increase of survival probability with increasing blanking duration (Fig. 6A, third panel) and the corresponding decrease of switch probability across the blank (Fig. 6B, third panel). In addition, there are fewer (up to 25.2%) presentations during which a dominant percept cannot be identified and these situations occur mainly at short blanking durations (Fig. 6C). This is in line with the experimental data, where no dominant percept was reported in up to 18.8% of the total experiment time. This happened primarily at short blanking durations, possibly due to the short time between presented stimuli. The model makes an important further prediction, namely that survival probability should decrease with stronger input. Our experimental data (Fig. 6, right column)—at least qualitatively—confirms this prediction.

Discussion

In this paper, we argue that rivalry can be understood as the result of a competition, just like attention can be understood as competition with priority control. WTA networks have been suggested as models of attention[52,61,62] and combining two WTA circuits together results in networks that have memory states.[48] Here, we demonstrated that a WTA model with state dependence replicates all key features of rivalry. In contrast, we found that simpler models without a memory state were unable to reproduce key aspects of rivalry. In particular, stateless models were unable to reproduce the phenomena of blanking. We conclude that memory plays an important role in competitive processes. Our model provides a first approach to how rivalry, attention,

and memory can be integrated into a single neuronally motivated model.

Limits of the present model

The present model was constructed to reproduce the key aspects common to nearly all forms of rivalry. As such, the model does not reproduce each and every aspect of any given rivalry experiment. In particular, we did not explicitly model time constants in a quantitative fashion. Furthermore, the input strength will depend on experimental details (as does the definition of what constitutes input strength in the first place[63]), and in the case of blanking, the survival probabilities in some cases can take far lower values than those found in our simulations and experiments. Models of a specific rivalry phenomenon would then have to include the upstream sensory circuits that realistically represent the input, where we here just made the reasonable but simplifying assumption that log stimulus contrast maps linearly to input currents. A specific model would also need to include the motor representation of the effector to report the precept[64,65] and include a notion of the rivalry stimulus' spatial extent to capture the spatial dynamics of dominance transitions.[66,67] Unlike in the experimental data and in contrast to their excess in models 1 and 2, extremely short dominance durations are absent for the double WTA model. To some extent this is a tradeoff between switching and memory, and to some extent it is a consequence of our definition of dominance (twice the other activity). This criterion was chosen to mimic the notion of (near) exclusive dominance in the experimental condition, and indeed, periods of no report were similar in frequency in model 3 and the experiment. While this is clearly a limitation of the present model, which has no natural mapping of its time axis to experimental time, a more detailed downstream readout and modeling the spatial distribution of dominance at any given point in time, will presumably allow relaxation of this criterion.

All these restrictions notwithstanding, with the double WTA we succeeded in modeling key properties that are common to all forms of rivalry in a single model: leptokurtic distributions, Levelt's propositions, and blanking.

Other modeling approaches for rivalry

Many attempts have been taken to model rivalry, capturing specific aspects. Many models of rivalry replicate the leptokurtic distribution of dominance durations.[54,68–73] Some of them also account for Levelt's second proposition, even though all of them only tested its original version fixed at highest input strength and did not investigate behavior at other fixed input strengths.[54,68,71] Levelt's fourth proposition has also been simulated by some of the existing models,[68,72] but how switch rate behaves under asymmetrical input has not been reported. The stabilizing effect of stimulus removal has been replicated over a large range of blanking and presentation durations by Noest *et al.*[74] and has been refined with a multi-timescale extension by Brascamp *et al.*[75] to cover their experimental findings. Still, this model is specifically designed to account for blanking behavior and percept choice at stimulus onset, leaving Levelt's propositions unaddressed. Wilson[54] extended his network to incorporate memory and thereby replicated the basic stabilizing effect of blanking, but leaves the functional relation between blank duration and survival probability unaddressed. The model of Gigante *et al.*[73] also accounts for blanking but leaves Levelt's propositions unaddressed. Thus, most networks perform well in replicating some of the key hallmarks of rivalry, but rarely are all of them addressed in a single framework. Only very few networks target all characteristics of rivalry and if they do, the whole range of input strength is not investigated. Hence, the double WTA network we presented here is the first to address Levelt's propositions, as well as blanking, for a wide range of input strengths.

Rivalry and memory

The key motivation for using the double WTA network is the memory state implicitly modeled by its dynamics. Consequently, only this network was able to reproduce the phenomenon of blanking. Of particular note is that this form of memory resulted in replicating Levelt's propositions. For a long time, rivalry was considered a memoryless process and thus, successive dominance durations were assumed to be independent and the timing of switches are unpredictable.[35,76,77] Recently, this notion has been challenged experimentally, both on short timescales of a few transitions[40,78] up to long-term fluctuations.[79] In addition, some physiological measures that have been tied to rivalry, such as eye position,[25,80,81] (micro-)saccade frequency,[82,83] eye blinks, and pupil size,[84,85] can also be used as

predictors of subsequent dominance,[81,84] again arguing for some information about subsequent states being available and thus against a memory-free process.

Rivalry and attention

As discussed above, we note that attention and rivalry are conceptually similar competitive processes. Many of the markers of rivalry, including eye position, saccades, and pupil size are also markers of attentional processes. In addition, attention and rivalry are also related behaviorally. Already von Helmholtz noted in his discussion of Schröder's[26] staircase and related multi-stable figures that he could volitionally switch his percept.[17] Similarly, for binocular rivalry, von Helmholtz[17] states that he could exert attentional control to keep one pattern dominant—an "arbitrary" amount of time for a simple line stimulus and by performing a task (e.g., counting) with the respective percept for more complex patterns. Recent research agrees with this notion: although transitions in rivalry seem to be spontaneous, some degree of volitional control can be exerted[86,87] and usefulness for the task can increase the dominance of the corresponding perceptual state.[88] Attention to a stimulus speeds up rivalry switching[89,90] and if attention is withdrawn from a stimulus, it has a stabilizing effect similar to stimulus removal; that is, rivalry is essentially abolished.[91,92] While the relation between attention and rivalry is a topic of intense research, these phenomena have so far been regarded as separate. In contrast, we here propose that both constitute a form of competition. This study is a first attempt to include memory and rivalry in a common model that has also been used to model attention. We expect that the present model can be extended to explicitly model the interactions between rivalry, memory, and attention. In summary, our model suggests that competition—with or without priority control—is a fundamental principle that links seemingly distinct phenomena.

Acknowledgments

This work was supported by the German Research Foundation (DFG) through Grant No. EI 852/3 (W.E.), the IRTG 1901 "The brain in action" (W.E., S.M.), SFB/TRR135 (W.E.), the GRK 1589/1 (G.G.), the Gustavus and Louise Pfeiffer Research Foundation (U.R.), and by an NIH Conte Center Grant Grant No. 5P50MH094258 (U.R.). S.M., G.G., U.R. and W.E. conceived the study; S.M., G.G., and U.R. conducted the modeling; S.M. and D.W. designed, conducted, and analyzed the experimental data; and S.M., U.R., and W.E. wrote the paper.

Conflicts of interest

The authors declare no conflict of interest.

Supporting Information

Additional supporting information may be found in the online version of this article.

Figure S1. Levelt's second proposition.

Table S1. Network parameters

References

1. James, W. 1890. *The Principles of Psychology*. New York: Holt.
2. Desimone, R. & J. Duncan. 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18:** 193–222.
3. Duncan, J. 1996. "Cooperating brain systems in selective perception and action." In *Attention and Performance XVI: Information Integration in Perception and Communication.* T. Inui & J.L. McClelland, Eds.: 549–578. Cambridge, MA: MIT Press.
4. Reynolds, J.H., L. Chelazzi & R. Desimone. 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* **19:** 1736–1753.
5. Duncan, J. 2006. EPS Mid-Career Award 2004: brain mechanisms of attention. *Q. J. Exp. Psychol.* **59:** 2–27.
6. Chelazzi, L., J. Duncan, E.K. Miller, *et al.* 1998. Responses of neurons in inferior temporal cortex during memory-guided visual search. *J. Neurophysiol.* **80:** 2918–2940.
7. Olivers, C.N.L., J. Peters, R. Houtkamp, *et al.* 2011. Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* **15:** 327–334
8. Zelinsky, G.J. & J.W. Bisley. 2015. The what, where, and why of priority maps and their interactions with visual working memory. *Ann. N.Y. Acad. Sci.* doi: 10.1111/nyas.12606.
9. Downing, P.E. 2000. Interactions between visual working memory and selective attention. *Psychol. Sci.* **11:** 467–473.
10. Awh, E., E.K. Vogel & S.-H. Oh. 2006. Interactions between attention and working memory. *Neuroscience* **139:** 201–208.
11. Theeuwes, J., A. Belopolsky & C.N.L. Olivers. 2009. Interactions between working memory, attention and eye movements. *Acta Psychol. (Amst.)* **132:** 106–114.
12. Hollingworth, A. & S. Hwang. 2013. The relationship between visual working memory and attention: retention of precise colour information in the absence of effects on perceptual selection. *Philos. Trans. R. Soc. B Biol. Sci.* **368:** 20130061.
13. Bundesen, C. 1990. A theory of visual attention. *Psychol. Rev.* **97:** 523–547.
14. Bundesen, C., T. Habekost & S. Kyllingsbæk. 2005. A neural theory of visual attention: bridging cognition and neurophysiology. *Psychol. Rev.* **112:** 291–328.

15. Broadbent, D.E. 1971. *Decision and Stress*. London: Academic Press.
16. Kyllingsbæk, S., B. Markussen & C. Bundesen. 2012. Testing a Poisson counter model for visual identification of briefly presented, mutually confusable single stimuli in pure accuracy tasks. *J. Exp. Psychol. Hum. Percept. Perform.* **38:** 628–642.
17. Von Helmholtz, H. 1867. *Handbuch der physiologischen Optik*. Leipzig, Germany: Voss.
18. Wertheimer, M. 1938. "Laws of organization in perceptual forms." In *A Source Book of Gestalt Psychology*. W. Ellis, Ed.: 71–88. London: Routledge & Kegan Paul.
19. Bülthoff, H.H. & A.L. Yuille. 1996. "A Bayesian framework for the integration of visual modules." In *Attention and Performance XVI: Information Integration in Perception and Communication*. T. Inui & J.L. McClelland, Eds.: 49–70. Cambridge, MA: MIT Press.
20. Elder, J.H. & R.M. Goldberg. 2002. Ecological statistics of Gestalt laws for the perceptual organization of contours. *J. Vis.* **2:** 324–353.
21. Freeman, W.T. 1994. The generic viewpoint assumption in a framework for visual perception. *Nature* **368:** 542–545.
22. Stocker, A.A. & E.P. Simoncelli. 2006. Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* **9:** 578–585.
23. Knill, D.C. & A. Pouget. 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27:** 712–719.
24. Körding, K.P. & D.M. Wolpert. 2004. Bayesian integration in sensorimotor learning. *Nature* **427:** 244–247.
25. Necker, L.A. 1832. LXI. Observations on some remarkable optical phenomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *London Edinburgh Philos. Mag. J. Sci.* **3:** 329–337.
26. Schröder, H. 1858. Ueber eine optische Inversion bei Betrachtung verkehrter, durch optische Vorrichtung entworfener, physischer Bilder. *Ann. Phys.* **181:** 298–311.
27. Rubin, E. 1915. *Synsoplevede figurer: studier i psykologisk analyse. 1. del*. Koebenhavn, Denmark: Gyldendalske Boghandel, Nordisk Forlag.
28. Boring, E.G. 1930. A new ambiguous figure. *Am. J. Psychol.* **42:** 444–445.
29. Adelson, E.H. & J.A. Movshon. 1982. Phenomenal coherence of moving visual patterns. *Nature* **300:** 523–525.
30. Hupé, J.-M. & N. Rubin. 2003. The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids. *Vis. Res.* **43:** 531–548.
31. Wheatstone, C. 1838. Contributions to the physiology of vision. Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philos. Trans. R. Soc. Lond.* **128:** 371–394.
32. Blake, R. 2001. A primer on binocular rivalry, including current controversies. *Brain Mind* **2:** 5–38.
33. Klink, P.C., R. Van Ee & R.J.A. van Wezel. 2008. General validity of Levelt's propositions reveals common computational mechanisms for visual rivalry. *PLoS One* **3:** e3473.
34. Meng, M. & F. Tong. 2004. Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J. Vis.* **4:** 539–551.
35. Levelt, W.J.M. 1967. Note on the distribution of dominance times in binocular rivalry. *Br. J. Psychol.* **58:** 143–145.
36. Levelt, W.J.M. 1968. *On Binocular Rivalry*. The Hague, Netherlands: Mouton.
37. Orbach, J., D. Ehrlich & A.H. Helen. 1963. Reversibility of the Necker cube. I. An examination of the concept of "satiation of orientation." *Percept. Mot. Skills* **17:** 439–458.
38. Leopold, D.A., M. Wilke, A. Maier, *et al.* 2002. Stable perception of visually ambiguous patterns. *Nat. Neurosci.* **5:** 605–609.
39. Schneider, W.X. 2013. Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Philos. Trans. R. Soc. B Biol. Sci.* **368:** 20130060.
40. Naber, M., G. Gruenhage & W. Einhäuser. 2010. Tri-stable stimuli reveal interactions among subsequent percepts: rivalry is biased by perceptual history. *Vision Res.* **50:** 818–828.
41. Carter, O. & P. Cavanagh. 2007. Onset rivalry: brief presentation isolates an early independent phase of perceptual competition. *PLoS One* **2:** e343.
42. Yuille, A.L. & D. Geiger. 2003. "Winner-take-all networks." In *The Handbook of Brain Theory and Neural Networks*. M.A. Arbib, Ed.: 1228–1231. Cambridge, MA: MIT Press.
43. Rutishauser, U., R.J. Douglas & J.-J. Slotine. 2011. Collective stability of networks of winner-take-all circuits. *Neural Comput.* **23:** 735–773.
44. Douglas, R.J., C. Koch, M. Mahowald, *et al.* 1995. Recurrent excitation in neocortical circuits. *Science* **269:** 981–985.
45. Douglas, R.J. & K.A.C. Martin. 2004. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* **27:** 419–451.
46. Binzegger, T., R.J. Douglas & K.A.C. Martin. 2004. A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* **24:** 8441–8453.
47. Neftci, E.O., B. Toth, G. Indiveri, *et al.* 2012. Dynamic state and parameter estimation applied to neuromorphic systems. *Neural Comput.* **24:** 1669–1694.
48. Rutishauser, U. & R.J. Douglas. 2009. State-dependent computation using coupled recurrent networks. *Neural Comput.* **21:** 478–509.
49. Itti, L. & C. Koch. 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis. Res.* **40:** 1489–1506.
50. Tsotsos, J.K., S.M. Culhane, W.Y.K. Wai, *et al.* 1995. Modeling visual attention via selective tuning. *Artif. Intell.* **78:** 507–545.
51. Lee, D.K., L. Itti, C. Koch, *et al.* 1999. Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* **2:** 375–381.
52. Hahnloser, R., R.J. Douglas, M. Mahowald, *et al.* 1999. Feedback interactions between neuronal pointers and maps for attentional processing. *Nat. Neurosci.* **2:** 746–752.
53. Matsuoka, K. 1984. The dynamic model of binocular rivalry. *Biol. Cybern.* **49:** 201–208.
54. Wilson, H.R. 2007. Minimal physiological conditions for binocular rivalry and rivalry memory. *Vis. Res.* **47:** 2741–2750.
55. Watson, A.B. & D.G. Pelli. 1983. QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* **33:** 113–20.

56. Brascamp, J.W., R. van Ee, A.J. Noest, *et al.* 2006. The time course of binocular rivalry reveals a fundamental role of noise. *J. Vis.* **6:** 1244–1256.
57. Brainard, D.H. 1997. The psychophysics toolbox. *Spat. Vis.* **10:** 433–436.
58. Pelli, D.G. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10:** 437–442.
59. Aafjes, M., J.E. Hueting & P. Visser. 1966. Individual and interindividual differences in binocular retinal rivalry in man. *Psychophysiology* **3:** 18–22.
60. Moreno-Bote, R., A. Shpiro, J. Rinzel, *et al.* 2010. Alternation rate in perceptual bistability is maximal at and symmetric around equi-dominance. *J. Vis.* **10:** 1.
61. Standage, D.I., T.P. Trappenberg & R.M. Klein. 2005. Modelling divided visual attention with a winner-take-all network. *Neural Netw.* **18:** 620–627.
62. Koch, C. & S. Ullman. 1985. Shifts in selective visual attention: towards the underlying neural circuitry. *Hum. Neurobiol.* **4:** 219–227.
63. Kang, M.-S. 2009. Size matters: a study of binocular rivalry dynamics. *J. Vis.* **9:** 17.1–11.
64. Wohlschläger, A. 2000. Visual motion priming by invisible actions. *Vis. Res.* **40:** 925–930.
65. Beets, I.A.M., F. Rösler, D.Y.P. Henriques, *et al.* 2010. Online action-to-perception transfer: only percept-dependent action affects perception. *Vis. Res.* **50:** 2633–2641.
66. Wilson, H.R., R. Blake & S.-H. Lee. 2001. Dynamics of travelling waves in visual perception. *Nature* **412:** 907–910.
67. Naber, M., O. Carter & F.A.J. Verstraten. 2009. Suppression wave dynamics: visual field anisotropies and inducer strength. *Vis. Res.* **49:** 1805–1813.
68. Laing, C.R. & C.C. Chow. 2002. A spiking neuron model for binocular rivalry. *J. Comput. Neurosci.* **12:** 39–53.
69. Wilson, H.R. 2003. Computational evidence for a rivalry hierarchy in vision. *Proc. Natl. Acad. Sci. USA* **100:** 14499–14503.
70. Lago-Fernández, L.F. & G. Deco. 2002. A model of binocular rivalry based on competition in IT. *Neurocomputing* **44:** 503–507.
71. Freeman, A.W. 2005. Multistage model for binocular rivalry. *J. Neurophysiol.* **94:** 4412–4420.
72. Moreno-Bote, R., J. Rinzel & N. Rubin. 2007. Noise-induced alternations in an attractor network model of perceptual bistability. *J. Neurophysiol.* **98:** 1125–1139.
73. Gigante, G., M. Mattia, J. Braun, *et al.* 2009. Bistable perception modeled as competing stochastic integrations at two levels. *PLoS Comput. Biol.* **5:** e1000430.
74. Noest, A.J., R. Van Ee, M.M. Nijs, *et al.* 2007. Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *J. Vis.* **7:** 10.
75. Brascamp, J.W., J. Pearson, R. Blake, *et al.* 2009. Intermittent ambiguous stimuli: implicit memory causes periodic perceptual alternations. *J. Vis.* **9:** 3.1–23.
76. Blake, R., R. Fox & C. McIntyre. 1971. Stochastic properties of stabilized-image binocular rivalry alternations. *J. Exp. Psychol.* **88:** 327–332.
77. Fox, R. & J. Herrmann. 1967. Stochastic properties of binocular rivalry alternations. *Atten. Percept. Psychophys.* **2:** 432–436.
78. Van Ee, R. 2009. Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: evidence from serial correlations in perceptual bistability. *J. Opt. Soc. Am. A* **26:** 2612–2622.
79. Suzuki, S. & M. Grabowecky. 2002. Evidence for perceptual "trapping" and adaptation in multistable binocular rivalry. *Neuron* **36:** 143–157.
80. Einhäuser, W., K.A.C. Martin & P. König. 2004. Are switches in perception of the Necker cube related to eye position? *Eur. J. Neurosci.* **20:** 2811–2818.
81. Kietzmann, T.C., S. Geuter & P. König. 2011. Overt visual attention as a causal factor of perceptual awareness. *PLoS One* **6:** e22614.
82. Van Dam, L.C.J. & R. van Ee. 2006. The role of saccades in exerting voluntary control in perceptual and binocular rivalry. *Vis. Res.* **46:** 787–799.
83. Van Dam, L.C.J. & R. van Ee. 2005. The role of (micro) saccades and blinks in perceptual bi-stability from slant rivalry. *Vis. Res.* **45:** 2417–2435.
84. Einhäuser, W., J. Stout, C. Koch, *et al.* 2008. Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proc. Natl. Acad. Sci. USA* **105:** 1704–1709.
85. Hupé, J.-M., C. Lamirel & J. Lorenceau. 2009. Pupil dynamics during bistable motion perception. *J. Vis.* **9:** 10.
86. Van Ee, R., L.C.J. Van Dam & G.J. Brouwer. 2005. Voluntary control and the dynamics of perceptual bi-stability. *Vis. Res.* **45:** 41–55.
87. Kornmeier, J., C.M. Hein & M. Bach. 2009. Multistable perception: when bottom-up and top-down coincide. *Brain Cogn.* **69:** 138–147.
88. Chopin, A. & P. Mamassian. 2011. Usefulness influences visual appearance in motion transparency depth rivalry. *J. Vis.* **11:** 18.
89. Paffen, C.L.E., D. Alais & F.A.J. Verstraten. 2006. Attention speeds binocular rivalry. *Psychol. Sci.* **17:** 752–756.
90. Paffen, C.L.E. & D. Alais. 2011. Attentional modulation of binocular rivalry. *Front. Hum. Neurosci.* **5:** 105.
91. Brascamp, J.W. & R. Blake. 2012. Inattention abolishes binocular rivalry: perceptual evidence. *Psychol. Sci.* **23:** 1159–1167.
92. Ling, S. & R. Blake. 2012. Normalization regulates competition for visual awareness. *Neuron* **75:** 531–540.

Online Methods

Single WTA (model 1 and 2)

The dynamics of the single WTA models (model 1 and model 2) are implemented using the following differential equations:

$$\begin{aligned} \tau \dot{u}_j + R u_j &= f(I_j + \alpha u_j - \beta_1 u_3 - T_j - \gamma A) \\ \tau_{ad} \dot{A}_j &= -A_j + u_j \\ \tau \dot{u}_3 + R u_3 &= f(\beta_2 \sum_{m=1}^{2} u_m - T_3) \end{aligned} \qquad \text{for j = 1,2}$$

Where τ is the integration time constant, R is the load or conductance and is set R = 1. $I_j(t)$ is the external input to unit j, α, β_1 and β_2 are positive (constant) parameters. γ is the self-adaptation gain and τ_{ad} is the adaptation time constant which is much slower than τ. The thresholds T_j are positive, equal for all units and constant in this circuit. f(x) is a non-saturating rectification non-linearity f(x) = max(0,x). In the single WTA without adaptation (model 1), γ is set to zero. All other values are identical in model 1 and 2 and given in table 1. Inputs applied to u1 and u2 have a Gaussian distribution, whose mean ("input strength") is fixed in each simulation and whose standard deviation is set to 1 independent of input strength. For models 1 and 2, input strengths in a range of 6.0 to 7.0 are tested in steps of 0.1.

Double WTA (model 3)

For model 3, the two coupled WTA circuits I and P are described by the following coupled differential equations:

$$\begin{aligned} P + \tau \frac{dP}{dt} &= f(W * P + X1 * I - \vec{t}) \\ I + \tau \frac{dI}{dt} &= f(W * I + X2 * P + Inp - \vec{t}_I) \end{aligned}$$

Upper case letters indicate matrices and lower case letters with arrows indicate vectors. The matrices describing the connection strengths within and between the circuits are:

$$W = \begin{bmatrix} \alpha & 0 & -\beta_1 \\ 0 & \alpha & -\beta_1 \\ \beta_2 & \beta_2 & 0 \end{bmatrix},\ X1 = \begin{bmatrix} \varphi & 0 & 0 \\ 0 & \varphi & 0 \\ 0 & 0 & 0 \end{bmatrix},\ X2 = \begin{bmatrix} 0 & \varphi & 0 \\ \varphi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Here, τ is the same time constant as in the single WTA, W is the weight matrix of all single WTA circuits (note: the parameters differ between single and double WTA). X1 contains feedforward weights from circuit I to circuit P while X2 contains the weights of feedforward input from P to I. Parameters used for all simulations are given in table 1.

External inputs are applied to i1 and i2 and follow a Gaussian distribution as in the case of models 1 and 2. Input strengths reach from 5.0 to 6.0 in steps of 0.1.

Simulations

Numerical integration was done using Euler integration with an integration time constant of δ=0.001. After every integration step, the nonlinear rectification function f(x) = max(x,0) was applied to each circuit.

The parameters were chosen such that stability is guaranteed. The criteria for stability, obtained by contraction analysis, are[38]:

$$0 < \alpha < 2\sqrt{\beta_1\beta_2}$$
$$0 < \beta_2$$
$$0 < \beta_1\beta_2 < 1$$
$$\gamma < 2\sqrt{\beta_1\beta_2} - \alpha$$

Changes in these parameter values can modulate median dominance duration and the spread of the distribution. Since the focus of the present paper is on the effect of varying input strength, rather than on the parameters governing network dynamics, we chose a fixed parameter set for each condition (single and double WTA) such that the systems remain stable and show alternating behavior for a range of input strengths. The chosen parameter values (table 1) are in the range typically used [38,43] and result in a high gain G ($G = \frac{1}{1+\beta_1\beta_2-\alpha}$).

Noise is updated independently for each input unit every 100 integration steps. The exact

noise update frequency is not critical, given that it is substantially larger than the integration time constant δ and substantially smaller than τ [43]. As long as these criteria are fulfilled, the value of the time constant τ is also uncritical and was here chosen to be 0.2 for all simulations for reasons of numerical efficiency. A natural time unit for all simulations is obtained by factoring out the arbitrary δ from the number of integration steps and is thereby given in units of $1/\delta$ (i.e., 1000 integration steps). All times reported for simulations are given in this unit. Each simulation was run for 1,500,000 integration steps; that is, for 1500 units of time and repeated five times with different random values for the noise.

Table 1: Network parameters

Single WTA		Double WTA	
Parameter	Value	Parameter	Value
τ	0.2	τ	0.2
α	1.2	α	1.4
β_1	3	β_1	3
β_2	0.2	β_2	0.2
γ	0.5 (model 2) 0 (model 1)	φ	0.25
τ_{ad}	200	t	[0.5 0.5 0.5]
$T_{1,2}$	6	t_I	[6 6 0.5]
T_3	0.5		

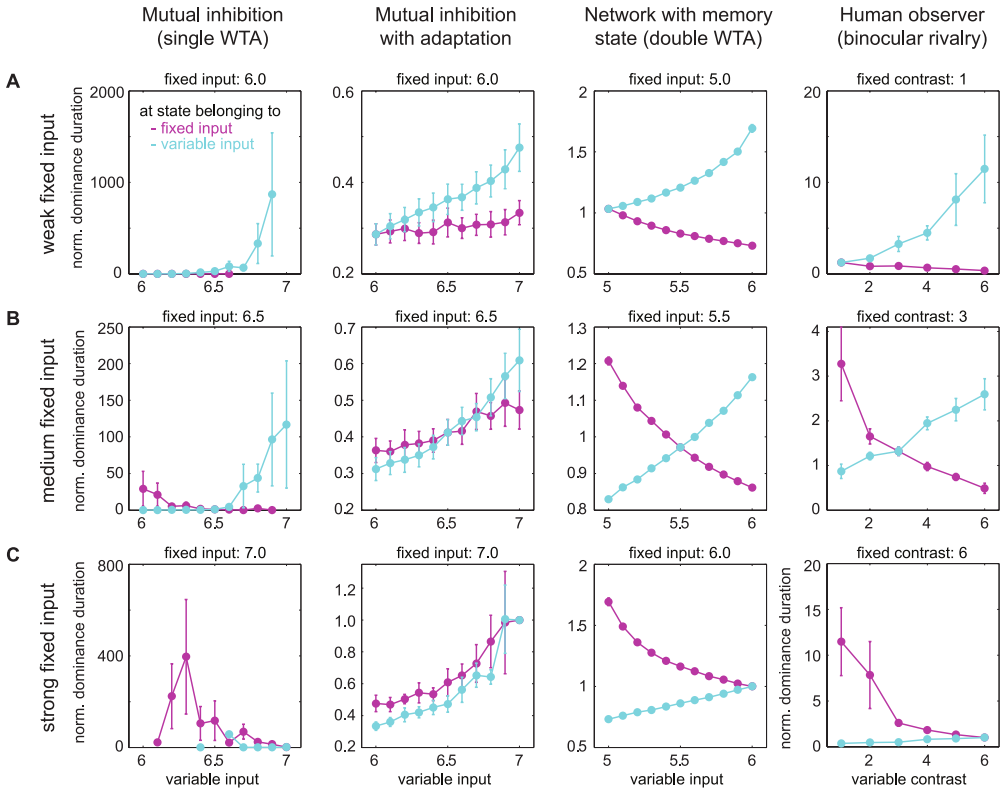

Figure S1. Levelt's second proposition. Different representation of the data presented in figure 4B. In each panel, the input to one eye/one input unit is fixed, the other is varied (x-axis). Normalized dominance durations in the fixed eye/input unit in magenta, for the varied eye/input unit in cyan. Columns denote different models and experimental data, rows different levels of fixed input: A) lowest, B) medium and C) highest input strength/contrast. All input strengths given in units of the respective currents, contrasts as contrast levels relative to the individual threshold as defined in the Methods section.

Journal of Vision (2014) 14(14):4, 1–18 http://www.journalofvision.org/content/14/14/4

Rapid serial processing of natural scenes: Color modulates detection but neither recognition nor the attentional blink

Svenja Marx Neurophysics, Philipps-University, Marburg, Germany

Onno Hansen-Goos Neurophysics, Philipps-University, Marburg, Germany

Michael Thrun Neurophysics, Philipps-University, Marburg, Germany

Wolfgang Einhäuser Neurophysics, Philipps-University, Marburg, Germany

The exact function of color vision for natural-scene perception has remained puzzling. In rapid serial visual presentation (RSVP) tasks, categorically defined targets (e.g., animals) are detected typically slightly better for color than for grayscale stimuli. Here we test the effect of color on animal detection, recognition, and the attentional blink. We present color and grayscale RSVP sequences with up to two target images (animals) embedded. In some conditions, we modify either the hue or the intensity of each pixel. We confirm a benefit of color over grayscale images for animal detection over a range of stimulus onset asynchronies (SOAs), with improved hit rates from 50 to 120 ms and overall improved performance from 90 to 120 ms. For stimuli in which the hue is inverted, performance is similar to grayscale for small SOAs and indistinguishable from original color only for large SOAs. For subordinate category discrimination, color provides no additional benefit. Color and grayscale sequences show an attentional blink, but differences between color and grayscale are fully explained by single-target differences, ruling out the possibility that the color benefit is purely attentional.

Introduction

The primate visual system is remarkably fast in grasping the "gist" of a complex natural scene (Biederman, 1972; Potter & Levy, 1969). Although the exact definition of what constitutes such a gist has remained elusive, Fei-Fei, Iyer, Koch, and Perona (2007) have provided a working—albeit somewhat circular—definition as the "contents of a glance." Experimental tests on the limits of perception within a glance frequently employ detection and/or recognition tasks. For example, observers are asked whether a scene contained a given high-level category (e.g., animal, means of transportation). When scenes are presented in isolation and without postmask, humans perform such tasks near ceiling for presentation durations as short as 20 ms (Thorpe, Fize, & Marlot, 1996). In that study, manual responses were given in under 300 ms, and the earliest category-dependent signal in the event-related potential (ERP) emerged as early as 150 ms after stimulus onset. In a later study that used a forced-choice saccade task, saccades had to be conducted to the hemifield where an animal had been briefly shown before, and some participants had reaction times as short as 120 ms (Kirchner & Thorpe, 2006). These findings are not restricted to animal targets, but are also valid for inanimate items (vehicles/no vehicles; VanRullen & Thorpe, 2001). Nonhuman primates show qualitatively similarly rapid categorization and are even somewhat faster than humans (Delorme, Richard, & Fabre-Thorpe, 2000; Fabre-Thorpe, Richard, & Thorpe, 1998).

Complementary to asking for possible neural implementations (Thorpe, Delorme, & VanRullen, 2001; Thorpe & Gautrais, 1997) of rapid scene processing, two questions arise on a behavioral level: First, which features are responsible for rapid recognition, and second, how does rapid recognition relate to attention processes? Wichmann, Drewes, Rosas, and Gegenfurtner (2010) addressed the former question and particularly the role of the power spectrum in rapid animal detection. They found that

Citation: Marx, S., Hansen-Goos, O., Thrun, M., & Einhäuser, W. (2014). Rapid serial processing of natural scenes: Color modulates detection but neither recognition nor the attentional blink. *Journal of Vision, 14*(14):4, 1–18, http://www.journalofvision.org/content/14/14/4, doi:10.1167/14.14.4.

doi: 10.1167/14.14.4 Received March 31, 2014; published December 16, 2014 ISSN 1534-7362

2.3 Rapid serial processing of natural scenes: Color modulates detection but neither recognition nor the attentional blink

a spectral cue eases animal detection without being causal. In another detailed analysis of the former question, Elder and Velisavljević (2009) investigated the role of several potential cues on visual processing in a rapid (30–120 ms) animal/no animal categorization task: two-dimensional boundary shape, texture, luminance, and color. They found that the fastest mechanisms relied on shape, while somewhat slower mechanisms integrated shapes with texture cues to become more robust. Color and luminance played virtually no role in this categorization task. Meng and Potter (2008) found similar results in an RSVP detection task with varying presentation durations (53, 107, 213, and 426 ms). Removing color information did not affect performance. In contrast, Delorme et al. (2010) investigated visual features for rapid (32-ms-presentation) animal categorization without postmask presentation in natural scenes and found a small but significant benefit of color in accuracy for responses later than 325 ms, while there was no benefit of color for the fastest responses. In addition to global image characteristics like luminance and color, they also tested the dependence of accuracy and reaction time on diagnostic animal features and target configuration. The most crucial features leading to high accuracy and speed turned out to be the presence of a typical animal posture and the area occupied by the animal (20%–30%). Wichmann, Braun, and Gegenfurtner (2006) reported an increase in performance of 2%–3% for colored as compared to grayscale pictures in a rapid animal/no animal categorization task. In monkeys and humans, color had a small but significant effect on reaction times when they had to detect food, but not when animals should be detected, and performance dropped slightly in some humans when color was removed (Delorme et al., 2000). The authors concluded that rapid identification may rely mainly on fast feed-forward processing of achromatic information in the magnocellular pathway.

In a rapid serial visual presentation (RSVP) paradigm, Yao and Einhäuser (2008) found again little effect of color on the detection of a single animal target among natural-scene distractors, though color boosted observers' confidence. In contrast, when participants were presented with two animal targets that belonged to different subordinate categories (species) within the same stream, the colored target was preferentially reported. This suggests that color, though having little effect on initial detection, plays a role for retrieval from memory. Not only retrieval from but also encoding into memory is influenced by color (Gegenfurtner & Rieger, 2000; Spence, Wong, Rusan, & Rastegar, 2006; Wichmann, Sharpe, & Gegenfurtner, 2002). Gegenfurtner and Rieger (2000) showed that color helps recognition in two ways, through adding a cue in coding at an early level and adding a cue in retrieval at a later stage. Thus they differentiated between the early, sensory influence and the later, cognitive influence of color. Although the benefit of color in early visual processing is small, it depends on the natural-scene content. If color is diagnostic for certain natural scenes (e.g., sea), it speeds up recognition without affecting accuracy (Oliva & Schyns, 2000; Rousselet, Joubert, & Fabre-Thorpe, 2005) and thus in these cases mediates rapid scene recognition. Nonetheless, the questions remain how the sensory influence of color develops over time, whether it affects detection of a superordinate category (e.g., animal) and the recognition of subordinate categories (e.g., animal species) alike, and whether or not color yields attentional benefits.

The role of attention in rapid visual processing of natural scenes has been the subject of many studies during recent years. When briefly peripherally flashed pictures had to be categorized into animals/no animals concurrently with an attentionally demanding central task, performance in both tasks did not drop as compared to single-task performance (Li, VanRullen, Koch, & Perona, 2002). Importantly, however, attention-demanding peripheral tasks like detecting a rotated *L* or *T* instead of detecting animals, led to a drop in performance, implying that animals can be detected even in the (near) absence of attention. This logic was later extended to specific subordinate classification tasks, such as gender discrimination (Reddy, Wilken, & Koch, 2004). Using a similar paradigm, Fei-Fei, VanRullen, Koch, and Perona (2005) found that grayscale pictures could also be processed very efficiently when attention was engaged elsewhere; furthermore, animal detection performance in the peripheral task was not impaired when a distractor image was shown in the periphery simultaneously with the target image at a location where either target or distractor could appear. This points out that early visual processing of natural scenes is not only nearly attention-free but also highly parallelized. This parallelization of early visual processing was also found by Rousselet, Fabre-Thorpe, and Thorpe (2002), who also used an animal/no animal categorization task showing either two pictures or one at the same time (left and/or right of central fixation) in a stream of pictures. Reaction times were the same in both conditions, and this was confirmed by category-related ERPs that emerged simultaneously (occipital: after 140 ms; frontal: after 160 ms) in both conditions and only differed after 190 ms.

When two items appear within close temporal succession in an RSVP stream, frequently an attentional blink (AB) is observed: A second target

(T2) is impaired when presented in a time window of 200–700 ms after a first target (T1). This decreased detection rate is usually absent if T2 immediately succeeds T1 ("lag-1-sparing"; Raymond, Shapiro, & Arnell, 1992). Initially, in these AB paradigms, artificial items were used (Chun & Potter, 1995; Raymond et al., 1992), but more recently a number of AB studies using natural scenes have been conducted. Evans and Treisman (2005) used the AB in their experiments 4 through 7 as a tool to test attentional effects on natural-scene processing. They presented a series of 12 natural scenes for 110 ms each, two of which contained targets. Here, target categories were animals and vehicles. When both targets had to be identified by giving a subordinate category, an AB was clearly measured and was more severe when targets were of different categories than when both targets were of the same category. There was also a subtle difference between categories, since animals were in general slightly better identified than vehicles. When both targets only had to be detected without identifying, the AB disappeared for targets of the same category and was only marginally present for sequences containing targets of different categories. Another study also found this dependency on stimulus category in an AB paradigm using natural scenes, where target categories were faces and watches (Einhäuser, Koch, & Makeig, 2007). Target identification was better and the AB was shorter for faces than for watches. Since the function of color vision in human beings and monkeys is frequently associated with attentional processes (Frey, Honey, & König, 2008; Maunsell & Treue, 2006; Motter, 1994; Zhang & Luck, 2009), the question arises whether color has an impact on the timing and depth of the AB.

To investigate the role of color in rapid visual processing and in particular its relation to attention, we conducted four RSVP experiments with animals as the target category. In the first experiment, observers in each trial had to report whether there were zero, one, or two animal targets in a 2-s stream, followed by a four-alternative forced-choice subordinate classification. Streams could be either colored or grayscale. This allowed us to replicate the small but frequently significant benefit of color for single-target processing and to characterize the dependence of subordinate classification on color and the modulation of the AB by color. In the second experiment, we asked whether the observed color benefits were a consequence of color being diagnostic for animals. To this end, we inverted the hue of each pixel (*roughly:* red-green, blue-yellow, etc.) while keeping saturation and luminance constant. In the third experiment, we tested whether the effects of color remained the same when stimulus presentation duration was decreased to 50 ms, using the same stimuli as in Experiment 2. And in the fourth experiment, we tested the dependence of color on six different presentation durations (which were also the stimulus onset asynchronies, SOAs) to test how the dependence on color develops over time.

Methods

In total we conducted four experiments. Experiment 1 targeted the effect of natural color in images on detection, recognition, and the time course of the attentional blink. Experiment 2 aimed at dissociating the effects of color that result from color's diagnosticity for animal images from those that result from other color-related effects. Experiment 3 investigated whether the results of Experiment 2 held for shorter SOAs. Experiment 4 analyzed the detection and recognition of single targets for a larger variety of SOAs.

Stimuli

A total of 480 animal target stimuli were used from the COREL data set: animals, vehicles, and distractors (http://vision.stanford.edu/resources_links.html; Li et al., 2002). For subordinate classification, animal images were subdivided into canine (including wolves, foxes, and dogs), feline (including tigers, pumas, and leopards), avian (including all kinds of birds), and ungulate (including horses, deer, cows, and goats), with 120 of each (Figure 1A). Distractor images were taken from the same database. In Experiments 1–3, the same subset of 360 target stimuli (90 per category) was used; in Experiment 4, all 480 were used. Stimuli were 384 × 256 pixels in size. We used four conditions, which we refer to as "original color," "color inverted," "grayscale," and "gray inverted." To modify stimuli, they were first transformed in the physiologically defined DKL color space (Derrington, Krauskopf, & Lennie, 1984). DKL color space is a three-dimensional space, in which the z-axis defines the luminance (for convenience, we map the minimal displayable luminance to −0.5 and the maximal displayable luminance to +0.5) and the other axes are spanned by the differential cone excitations: the difference between L and M cones (L − M axis) and the difference between the sum of L and M cones and S cones (S − (L + M) axis).

The original image (Figure 1A) was kept unchanged. For the grayscale condition (Figure 1B), saturation in the DKL space was set to 0 (i.e., each pixel was projected on the luminance axis). For the color-inverted condition (Figure 1C), the DKL space was rotated by 180°, which results in a mapping of each hue to its opponent hue without any change in saturation or

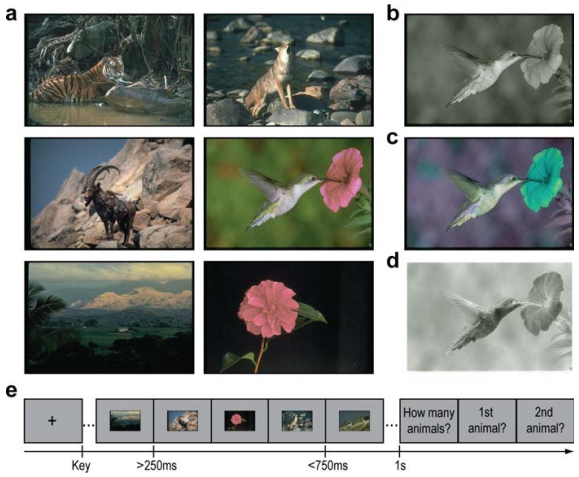

Figure 1. Stimuli and procedure. (a) Example stimuli of all four target categories (feline, canine, ungulate, avian) and two example distractor images. Image modifications: (b) grayscale, (c) color inverted, (d) gray inverted. (e) Procedure. Depicted times correspond to Experiments 1 and 2; in all experiments, targets occurred between serial positions 6 and 15.

luminance. Since it is not guaranteed that the modified image can be displayed within the screen's gamut, we applied the following procedure to keep luminance and inverted hue as unaffected as possible. After the hue inversion, we determined for each pixel the maximal chroma (distance from the luminance axis in DKL space) the screen could display for the given hue. If the chroma of the pixel was at or below this maximum, the pixel remained unchanged. If the chroma of the pixel was above this maximum, it was reduced to this maximally displayable value while keeping luminance and hue unchanged (i.e., both cardinal color axes were scaled by the same factor, while the luminance axis was not scaled). On average, 7.2% ± 9.9 pixels were affected by such a reduction, and on average the scaling factor for the axes was 0.98. That is, the reduction in chroma was small and affected only a small number of pixels. For the gray-inverted condition (Figure 1D), the luminance values were inverted as follows: Luminance was mapped to the interval [0, 1] (i.e., 0.5 was added to the luminance axis of the DKL space), the square root of the resulting values was subtracted from 1, that result was squared, and it was then mapped back to [–0.5, 0.5] by subtracting 0.5.

Procedure

In Experiments 1–3, observers viewed streams of 20 natural scenes that contained either no target, one target, or two targets (Figure 1E). In Experiment 4, all streams contained either no target or one target. All images in the stream (target and distractors) were subjected to the same color conditions. Observers were asked to fixate the center of the screen and press and release a button to start each trial. After viewing each stream, observers were first asked how many animals they had seen in the preceding stream. Then they were asked to choose the animal class (if they had responded "one") or classes in order (if they had responded "two"; Experiments 1–3) among the set of four options (feline, canine, avian, ungulate). The number of queries depended on the response, not on ground truth. That is, even if a detection was a false alarm, observers had to respond which category they had recognized, and they were not prompted for categorization if they had not detected any target.

In Experiment 1, the SOA was 100 ms and only grayscale and original-color conditions were used. For each color condition (grayscale, original), Experiment 1 included 240 streams with zero targets, 240 streams with one target, and 240 streams with two targets, 48 for each tested lag (one, two, three, four, and seven frames). This yielded a total of 1,440 ($2 \times 3 \times 5 \times 48$) trials. The order of trials was randomized and the experiment split in two sessions of about equal length. In Experiment 2, the SOA was also 100 ms, but each stream was presented in all four color conditions, with 120 streams of zero targets, 120 streams with one target, and 120 streams with two targets (all at lag-2) for each condition, again yielding a total of 1,440 ($4 \times 3 \times 120$) trials that were split in two sessions of about equal length. In Experiment 3, the SOA was 50 ms and the experiment was otherwise identical to Experiment 2. In Experiment 4, six different SOAs were used: the 50 and 100 ms of the previous experiments as well as 30, 60, 90, and 120 ms. Each stream was presented in all four color conditions with 80 streams of zero targets and 80 streams with one target per color condition and per SOA, yielding $2 \times 4 \times 6 \times 80 = 3{,}840$ trials. They were split in three sessions of about equal length. In Experiment 1, each of the 360 target stimuli was used twice per condition (in different streams of distractors); in Experiments 2 and 3, each of the 360 target stimuli was used once per condition; and in Experiment 4, each of the 480 target stimuli was used once per color condition.

Setup

The study was conducted in a dark and sound-isolated room. Stimuli were presented on a 19.7-in. EIZO Flex Scan F77S CRT monitor set to 1024 × 768 pixel resolution at 100 Hz, located at 73 cm distant from the observer, whose head was stabilized

with a chin rest and a forehead rest. The maximum luminance (white) was 66.0 cd/m^2, the minimum luminance (black) was 0.11 cd/m^2, and the CIE color coordinates of the monitor's guns (x, y) were (0.623, 0.344) for red, (0.287, 0.609) for green, and (0.151, 0.065) for blue. Stimuli spanned 11.6° × 7.8° on a gray background. Before each trial, a gray fixation screen with a black fixation cross was presented.

All stimulus preparation, presentation, and data analysis used Matlab (Mathworks, Natick, MA). For presentation, the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002), as retrieved from http://www.psychtoolbox.org, were used with Matlab.

Participants

Thirty-two volunteers participated in the study: eight (six female, 24.8 ± 3.3 years) in Experiment 1, eight (six female, 26.5 ± 2.8 years) in Experiment 2, eight in Experiment 3 (two female, 26.6 ± 5.8 years), and eight in Experiment 4 (six female, 25.1 ± 2.9 years). They were paid for participation, and experiments conformed with the Declaration of Helsinki and were approved by the local ethics committee (Ethikkommission FB04).

Analysis

Since the design for all experiments was "within-subject" for all variables of interest, all analyses treated observers as repeated measures. For analyses of more than one factor or more than two levels per factor, a repeated-measures ANOVA was used. For post hoc pairwise comparisons and for factors with two levels, paired t tests were used.

Two types of analysis have to be distinguished, hereafter "detection" and "recognition." Detection refers to the question whether the number of targets the observer reported corresponded to the number of targets present in the stream. We tested results for zero-target, one-target, and two-target trials separately, and refer to the relevant variables by standard signal-detection-theory terms.

For the first part of analysis in all experiments, we considered only single-target and no-target trials. For zero-target streams, we defined the report of any target (one or two) as a false alarm. For single-target streams there are two possible errors: the report of no target or the report of two targets. Although the latter was rare (see Appendix A for each individual's 3 × 3 matrix of all possible truth/response combinations), we performed the analysis for both definitions: at least one target reported and exactly one target reported. For the computation of d' (computed as the difference between the z-scored hit and false-alarm rates; Macmillan, 1993), we used the former definition. In Experiment 4, only zero or one target was possible, so that the hit and false-alarm rates are unambiguously defined.

In recognition, we asked whether the target was correctly identified according to the four available categories. Most analysis is based on "recognition given detection," that is, refers only to trials in which the target or targets were detected. In one-target streams for which two targets were reported, the target was counted as recognized if at least one of the two responses matched the target category. When analyzing recognition for two-target streams in which exactly one target was detected and T1 was of the same category as T2, it is impossible to infer from the response whether T1 or T2 was recognized (as both require the same response). For this particular analysis, we therefore excluded trials for which T1 and T2 were from the same category.

Results

Detection of single targets

For a first analysis, we consider zero-target and single-target trials (Figure 2; Tables 1 and 2). In Experiments 1 and 2, with an SOA of 100 ms, color sequences had more hits (Figure 2A; Appendix B) and fewer false alarms (Figure 2B; Appendix B) than grayscale sequences. The difference in hit rates was in the typically observed range, no matter whether hits were defined as response 1 or 2 in one-target trials (Experiment 1: 4.1% ± 3.8%; Experiment 2: 3.1% ± 2.3; all data mean ± SD) or as response of exactly 1 (Experiment 1: 4.8% ± 3.0%; Experiment 2: 4.0% ± 2.2%). Qualitatively, the same held for Experiment 3 (SOA: 50 ms), with a difference of 4.6% ± 4.4% (response > 0) or 5.1% ± 3.8% (response = 1) between color and grayscale. In Experiment 4, where hit rates are well defined, as only responses 0 and 1 are possible, color had higher hit rates than grayscale for all SOAs (Figure 2A, right). This qualitative benefit for color is also reflected in the separability (d'), which combines hits and false alarms into a measure of performance (Table 2; Figure 2C): The value of d' for color sequences is larger than for grayscale sequences across all conditions (Figure 3). In Experiments 2–4 we had two additional conditions: color inverted and gray inverted. The gray-inverted images have fewer hits, more false alarms, and consequently a smaller d' for all conditions tested (Figures 2 and 3, gray). The color-inverted condition (Figures 2 and 3, red) shows a more

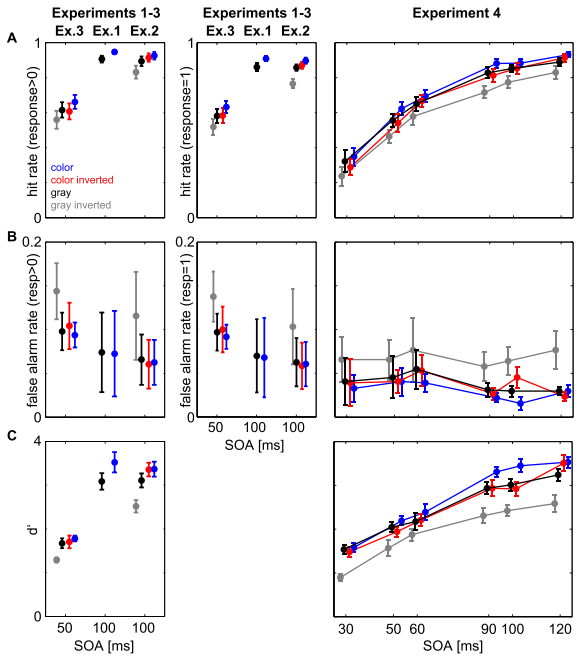

Figure 2. Detection in zero-target and single-target trials. (A) Hit rate for Experiments 1–3 (left, sorted by SOA) and the different SOAs of Experiment 4. Different colors code different conditions (blue: original color; red: color inverted; black: grayscale; gray: gray inverted). The left panel defines hits as any response to single-target trials (response 1 or 2), the middle panel as an exact response (response 1). For Experiment 4, there was no two-target option. (B) False alarms. Notation as in (A). Left panel: any false alarm (response 1 or 2); middle panel: single false alarm (response 1). (C) Value of d' as computed from z-scored hit and false-alarm rates of (A) and (B); the " >0" definition of hits and false alarms is used for this computation. Error bars are mean and standard error of the mean across observers.

mixed pattern: For small SOAs (Experiment 3 and the short SOAs of Experiment 4) it tends to be close to the grayscale condition, while for larger SOAs (Experiment 2 and the long SOAs of Experiment 4) it tends to be close to the color condition.

To quantify these effects statistically, for the experiments with more than two conditions (2–4) we first tested whether the factor color condition had an effect at all by means of a repeated-measures ANOVA (in Experiment 4 with SOA as an additional factor). For hits (in either definition), false alarms, and d' we find main effects of condition in all experiments (Table 1). In Experiment 4, we additionally find a main effect of SOA for hits and d' (though not for false alarms), but no interaction between condition and SOA (Table 1). This allowed us to perform post hoc tests for all experiments and each SOA level in Experiment 4, as to which color conditions differ from each other in terms of hits, false alarms, and/or d'. In the remainder of the main text we will focus on d'; hit and false-alarm data are analyzed in Appendix B.

When considering d' as a performance measure that combines over hits and false alarms and is thus insensitive to subjective criteria, the difference between color and grayscale sequences increases monotonically up to 100 ms (Figure 3) and becomes significant at 90 ms and above (Table 2). This indicates that there is a benefit induced by color that increases with increasing SOA, at least up to 100 ms.

To address whether the performance benefit derives from color being diagnostic for animal scenes, we included the color-inverted images in Experiments 2–4. For SOAs of 90 and 100 ms (Experiment 4), where color already excels over grayscale, the color-inverted sequences yield significantly worse performance than the original color sequences (Figure 3; Table 2). Only for the longest SOAs (100 ms in Experiment 2 and 120 ms in Experiment 4), color-inverted sequences yield (or trend to yield) better performance than grayscale and become indistinguishable from original color.

Performance in the gray-inverted condition is—with the exception of an SOA of 60 ms in Experiment 4, where it is indistinguishable from grayscale and color inverted—significantly worse than in any other condition (all $t(7) > 3.02$, all $p < 0.02$). As the target is clearly identifiable in these images if viewing time is infinite, the gray-inverted condition verifies that even at the largest SOAs tested, detection is not yet trivial (i.e., it is not equivalent to prolonged viewing).

In sum, the benefit of color for detection increases with increasing SOA (Figure 3, blue), but only at large SOAs can a similar benefit be observed for color-inverted images (Figure 3, red). This suggests that at short SOAs, the color benefit results from mechanisms that require the correct hue (e.g., the hue being diagnostic for target images), while for longer SOAs other mechanisms, which only require color contrasts to be intact, may come into play.

Color and grayscale targets both induce an attentional blink

In Experiment 1, we tested two-target streams at a variety of lags (1, 2, 3, 4, 7) between targets. When analyzing color and grayscale sequences separately (Appendix C), we find reduced performance for short lags, the attentional blink (AB). Performance is worst at lag 1; that is, we do not observe lag-1 sparing (Figure 4A). This absence of lag-1 sparing also holds when only trials with T1 and T2 from the same category are considered, ruling out the possibility that it results from the dissimilarity between categories.

	SOA	Effect of	Hits	False alarms	d'
Experiment 2	100 ms	Condition	**$F(3, 21) = 15.7$, $p = 1.38 \times 10^{-5}$ (response > 0)** **$F(3, 21) = 42.05$, $p = 4.7 \times 10^{-9}$ (response $= 1$)**	**$F(3, 21) = 4.57$, $p = 0.013$**	**$F(3, 21) = 35.77$, $p = 1.95 \times 10^{-8}$**
Experiment 3	50 ms	Condition	**$F(3, 21) = 11.6$, $p = 1.07 \times 10^{-4}$ (response > 0)** **$F(3, 21) = 17.68$, $p = 5.8 \times 10^{-6}$ (response $= 1$)**	**$F(3, 21) = 4.23$, $p = 0.017$**	**$F(3, 21) = 7.26$, $p = 0.0016$**
Experiment 4	All	SOA	**$F(5, 35) = 128.37$, $p < 10^{-20}$**	$F(5, 35) = 0.78$, $p = 0.57$	**$F(5, 35) = 87.10$, $p < 10^{-20}$**
		Condition	**$F(3, 21) = 69.89$, $p = 4.28 \times 10^{-11}$**	**$F(3, 21) = 6.65$, $p = 0.0025$**	**$F(3, 21) = 72.65$, $p = 2.96 \times 10^{-11}$**
		SOA × Condition	$F(15, 105) = 1.30$, $p = 0.22$	$F(15, 105) = 0.81$, $p = 0.66$	$F(15, 105) = 1.51$, $p = 0.11$

Table 1. ANOVAs for effect of color condition (Experiments 2–4) and SOA (Experiment 4) for hits, false alarms, and d'. In Experiments 2 and 3, two different definitions of hits are tested: response to a one-target trial of at least 1 ("> 0") or exactly 1 ("$= 1$"). Bold type indicates a significant effect.

If the detection of one target in a two-target stream were independent from the detection of the other, the probability of detecting both targets would equal the square of the single-target hit rate. Using this baseline, we find a significant AB at lag 2 in Experiment 2 (Figure 4B) and Experiment 3 (Figure 4C) for all color conditions (Appendix C). Hence, there is an attentional blink (without lag-1 sparing) for lags 1, 2, and 3 for any color condition and for short (50 ms) and long (100 ms) SOAs.

SOA	Gray inverted	Grayscale	Color inverted	Original color	Gray vs. color	Gray vs. color inverted	Color inverted vs. original color
Experiment 1							
100 ms		3.09 ± 0.54		3.52 ± 0.66	**$t(7) = 2.61$, $p = 0.035$**		
Experiment 2							
100 ms	2.52 ± 0.40	3.11 ± 0.46	3.36 ± 0.45	3.37 ± 0.49	**$t(7) = 2.81$, $p = 0.026$**	**$t(7) = 3.30$, $p = 0.013$**	$t(7) = 0.096$, $p = 0.93$
Experiment 3							
50 ms	1.29 ± 0.16	1.67 ± 0.32	1.70 ± 0.41	1.78 ± 0.19	$t(7) = 1.07$, $p = 0.32$	$t(7) = 0.23$, $p = 0.83$	$t(7) = 0.62$, $p = 0.56$
Experiment 4							
30 ms	0.90 ± 0.24	1.54 ± 0.31	1.48 ± 0.34	1.58 ± 0.31	$t(7) = 0.46$, $p = 0.66$	$t(7) = 0.39$, $p = 0.71$	$t(7) = 0.63$, $p = 0.55$
50 ms	1.56 ± 0.50	2.04 ± 0.33	1.94 ± 0.32	2.19 ± 0.29	$t(7) = 1.30$, $p = 0.23$	$t(7) = 0.80$, $p = 0.45$	$t(7) = 2.10$, $p = 0.074$
60 ms	1.88 ± 0.39	2.17 ± 0.56	2.21 ± 0.38	2.39 ± 0.50	$t(7) = 1.29$, $p = 0.24$	$t(7) = 0.22$, $p = 0.83$	$t(7) = 2.08$, $p = 0.076$
90 ms	2.31 ± 0.49	2.93 ± 0.38	2.93 ± 0.53	3.31 ± 0.33	**$t(7) = 3.34$, $p = 0.012$**	$t(7) = 0.0019$, $p = 1.0$	**$t(7) = 3.81$, $p = 0.0066$**
100 ms	2.42 ± 0.34	3.01 ± 0.40	2.92 ± 0.44	3.45 ± 0.44	**$t(7) = 3.16$, $p = 0.016$**	$t(7) = 0.67$, $p = 0.52$	**$t(7) = 4.54$, $p = 0.0027$**
120 ms	2.58 ± 0.53	3.24 ± 0.39	3.51 ± 0.53	3.53 ± 0.36	**$t(7) = 2.41$, $p = 0.047$**	$t(7) = 1.91$, $p = 0.098$	$t(7) = 0.16$, $p = 0.88$

Table 2. Values of d' and post hoc comparisons of interest. Bold type indicates a significant effect; all d' values are mean ± standard deviation.

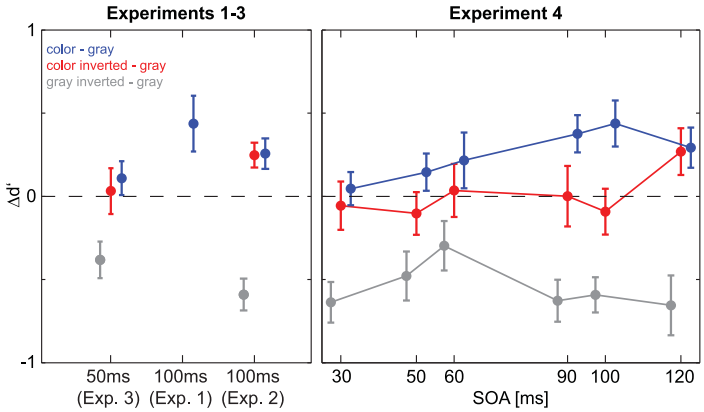

Figure 3. Value of d' difference to grayscale. Difference of the different color conditions to grayscale. Mean and standard error of the mean across observers.

Any effect of color on the attentional blink is explained by single-target performance alone

When testing the two-target detection rate at each lag, there seems an apparent effect of color: Detection performance is better for color conditions at all lags (all ts > 2.7, all ps < 0.03; Figure 4A). Similarly, there is a main effect of color condition on the two-target hit rate in Experiment 2, $F(3, 21) = 59.27$, $p = 2.0 \times 10^{-10}$ (Figure 4B), and in Experiment 3, $F(3, 21) = 3.96$, $p = 0.022$ (Figure 4C). This raises the question whether there is an attentional benefit of color or whether this effect can solely be explained by differences in single-target performance. To answer this question, we subtracted the baseline, defined by the individual's squared single-target hit rate in the respective color condition, from the plain two-target hit rate. For

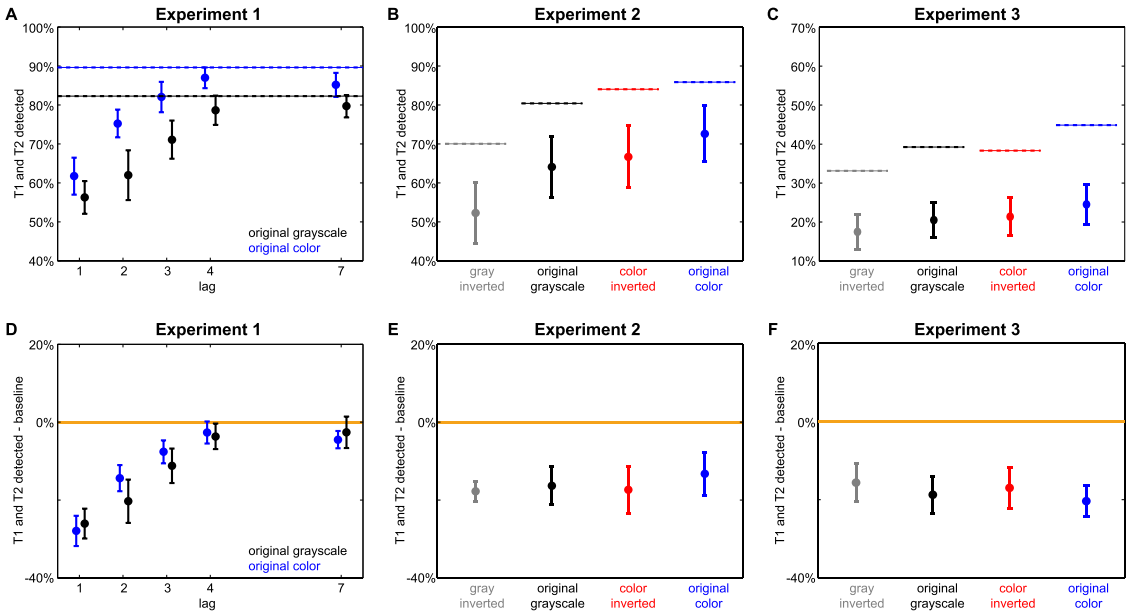

Figure 4. Attentional blink. (A–C) Detection rate for both targets in the two-target sequences; the dashed line indicates the baseline (squared detection rate of the single-target sequences) (A) at different lags in Experiment 1, (B) at lag 2 (200 ms) in Experiment 2, and (C) at lag 2 (100 ms) in Experiment 3. (D–F) Baseline-corrected detection rate in two-target sequences in (D) Experiment 1, (E) Experiment 2, and (F) Experiment 3.

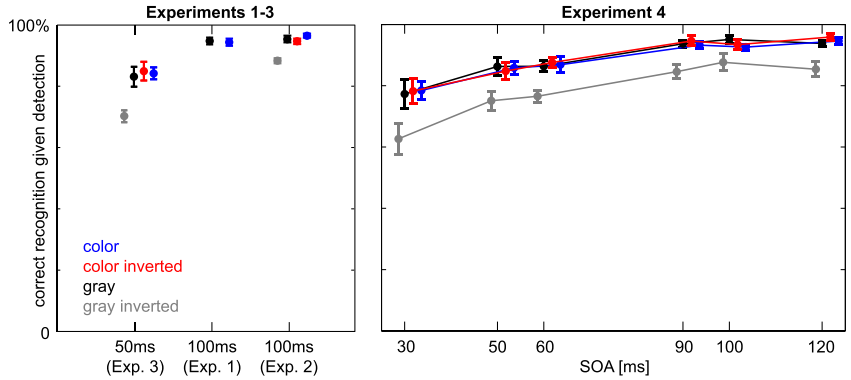

Figure 5. Recognition given detection. Number of single-target trials for which the correct category was reported divided by the number of single-target trials in which at least one target was reported (hits). Colors as in Figure 2; bars denote standard error of the mean across observers.

Experiment 1, we find that these baseline-corrected data do not show differences between color and grayscale at any lag (all ts < 0.69, all ps > 0.21; Figure 4D). Similarly, there is no main effect of color condition on two-target detection performance after baseline correction in Experiment 2, $F(3, 21) = 1.18$, $p = 0.34$ (Figure 4E) or Experiment 3, $F(3, 21) = 1.14$, $p = 0.36$ (Figure 4F). Consequently, while we find an attentional blink for two-target detection in any color condition, we do not observe any effect of color in addition to what single-target detection performance had predicted.

Color effects in single-target recognition are explained by detection performance

Besides the mere detection of animals in a sequence of distractors, we also tested the capability of observers to identify the subordinate category. Of all single-target trials in Experiment 1, the subordinate animal category is correctly identified in 85.8% ± 6.1% of grayscale and 89.3% ± 4.7% of color sequences, with a significant benefit of color, $t(7) = 3.30$, $p = 0.01$. Similarly, there is a significant effect of color condition on recognition in Experiment 2, $F(3, 21) = 56.95$, $p = 2.9 \times 10^{-10}$, and in Experiment 3, $F(3, 21) = 26.79$, $p = 2.26 \times 10^{-7}$. In Experiment 4, a 6 × 4 repeated-measures ANOVA reveals a main effect of SOA, $F(5, 35) = 265.31$, $p < 10^{-20}$, and of color condition, $F(3, 21) = 140.37$, $p = 4.72 \times 10^{-14}$, but no interaction, $F(15, 105) = 1.71$, $p = 0.059$. In line with the absence of an interaction, color condition has an effect on recognition at each SOA (Fs > 18.67, ps $< 3.90 \times 10^{-6}$), and in turn, SOAs have an effect on recognition in all color conditions (Fs > 118.3, ps $< 10^{-20}$).

This analysis considered recognition unconditionally; that is, it compared raw recognition rates independent of whether the target was at all detected. However, when only considering single-target trials in which the target is correctly detected (recognition given detection), subordinate recognition is indistinguishable between grayscale and color (Figure 5). In Experiment 1, 94.8% ± 3.2% of grayscale and 94.4% ± 3.4% of color sequences in which a target is correctly detected have the target also correctly identified, $t(7) = 0.48$, $p = 0.65$. For Experiments 2 and 3, there are still main effects of color condition on this recognition-given-detection performance—Experiment 2: $F(3, 21) = 27.4$, $p = 1.9 \times 10^{-7}$; Experiment 3: $F(3, 21) = 16.37$, $p = 1.02 \times 10^{-5}$—but this is solely explained by the difference between the gray-inverted category and all other categories (Appendix D; Figure 5). For the recognition-given-detection analysis in Experiment 4, there is a main effect of SOA, $F(5, 35) = 26.93$, $p = 4.46 \times 10^{-11}$, and of color, $F(3, 21) = 62.54$, $p = 1.22 \times 10^{-10}$, but no interaction, $F(15, 105) = 1.0$, $p = 0.46$. This main effect of color also results almost entirely from the difference between gray-inverted and all other conditions (Appendix D; Figure 5). In general, once animal detection has been successful, color has no additional effect on subordinate animal categorization. In contrast, the polarity of luminance has an effect.

Color effects in two-target recognition performance are explained by detection performance

For the data of Experiment 1, we tested whether lag and/or color has an effect on recognition performance in trials in which two targets are correctly detected by means of a two-way (2 color conditions × 5 lags)

repeated-measures ANOVA. We find a main effect of lag on the probability that both targets were correctly identified, $F(4, 28) = 4.62$, $p = 0.0055$, but no effect of color condition, $F(1, 28) = 0.0017$, $p = 0.97$, and no interaction, $F(4, 28) = 1.04$, $p = 0.40$. This shows that there is an attentional blink for recognition on top of that for detection, but no additional effect of color. Once detection partially fails, subsequent recognition is insensitive to color or the attentional blink, which also holds for Experiments 2 and 3 (Appendix E).

Discussion

The present study shows that color is beneficial for rapid scene perception ("color benefit"). Targets in rapidly presented sequences of images are slightly easier to detect when sequences are in color as compared to grayscale, which is qualitatively and quantitatively in line with several earlier reports (Delorme et al., 2010; Wichmann et al., 2006). We find that this sensory color benefit increases monotonically with presentation time up to about 100 ms. For short presentation times, the color benefit requires the hue to be intact, pointing out that color being diagnostic for images containing an animal may be the dominant effect driving the color benefit for short SOAs. For longer SOAs, hue-modified images tend to approach original-color performance, suggesting that a general benefit of color as such, possibly related to a segmentation process, comes into play. Color does not aid performance in naming subordinate animal categories, provided detection of the category "animal" was successful. Finally, color has no influence on the characteristic of the attentional blink beyond the effects explained by single-target trials alone. Together, these results suggest a preattentive, rather than attentional, source of the color benefit.

While some previous studies have also found an effect of color on performance in rapid detection (Delorme et al., 2010; Wichmann et al., 2006), others have not (Elder & Velisavljević, 2009; Meng & Potter, 2008). In an RSVP paradigm, Meng and Potter (2008) instructed participants about the target with short descriptions of the scene to expect and found no effect of color on the detection of the scene for a wide range of SOAs (53–426 ms) and for normal and impoverished viewing conditions alike. One possible explanation for the absence of an effect could be that color might be beneficial only for broad categories (like animals), not for more detailed descriptions, especially if they implicate a spatial relation (like the "businessmen at table" example in Meng & Potter, 2008). In contrast to the present study and Meng and Potter's, Elder and Velisavljević (2009)—who did not find an effect of color—did not use sequences of images but instead used masked presentations of isolated images. Depending on the exact design of the mask, the colored mask might be relatively more effective than a grayscale mask as compared to the difference between the temporally adjacent color or grayscale frames in RSVP. Whether the color benefit extends from animals to other categories, whether it depends on the richness of the instruction, whether it depends on whether the instructions imply spatial relations, and whether there is a fundamental difference between RSVP and isolated masked images are interesting issues for further research.

When detecting targets in complex backgrounds, the separation of figure from ground is an important role for segmentation processes. The interpretation that color facilitates figure–ground segmentation has been proposed by other studies, suggesting this mechanism as an early contribution of color to visual processing (Gegenfurtner & Rieger, 2000; Skiera, Petersen, Skalej, & Fahle, 2000). Wurm, Legge, Isenberg, and Luebker (1993) found that color improved accuracy in recognition of food targets irrespective of its diagnosticity of the target object, which points to a rapid low-level contribution of color to object recognition. This early contribution of color in visual processing and particularly in figure–ground segmentation has also been shown in an fMRI study where activity related to figure–ground segmentation in checkerboards by color, luminance, and motion stimuli was already found in the primary visual cortex (Skiera et al., 2000). For long SOAs, performance in color-inverted sequences trends more towards original-color performance than towards grayscale performance (Figure 3). Since segmentation in natural scenes benefits from chromatic boundaries (Hansen & Gegenfurtner, 2009), which are unaffected by our hue inversion, it is conceivable that for long SOAs, color aids detection by fostering segmentation independent of hue being diagnostic for the target category.

Although we focus on sensory aspects and find an effect of color on detection but not on recognition or attention, our results do not contradict the notion that color also plays a prominent role in later stages of visual processing. It has already been proposed that color aids visual processing not primarily during detection but in the later stages, when—for example—memory has to be accessed (Yao & Einhäuser, 2008). This prominent role of color in encoding and retrieval in recognition memory paradigms has been shown in several studies (Gegenfurtner & Rieger, 2000; Spence et al., 2006; Wichmann et al., 2002) and has typically exceeded the comparably subtle effect in rapid visual categorization tasks (Delorme et al., 2000, 2010; Wichmann et al., 2006).

Unlike the original characterization of the AB (Raymond et al., 1992), we do not observe lag-1 sparing in our Experiment 1 for color nor for grayscale sequences. Visser, Zuvic, Bischof, and Di Lollo (1999)

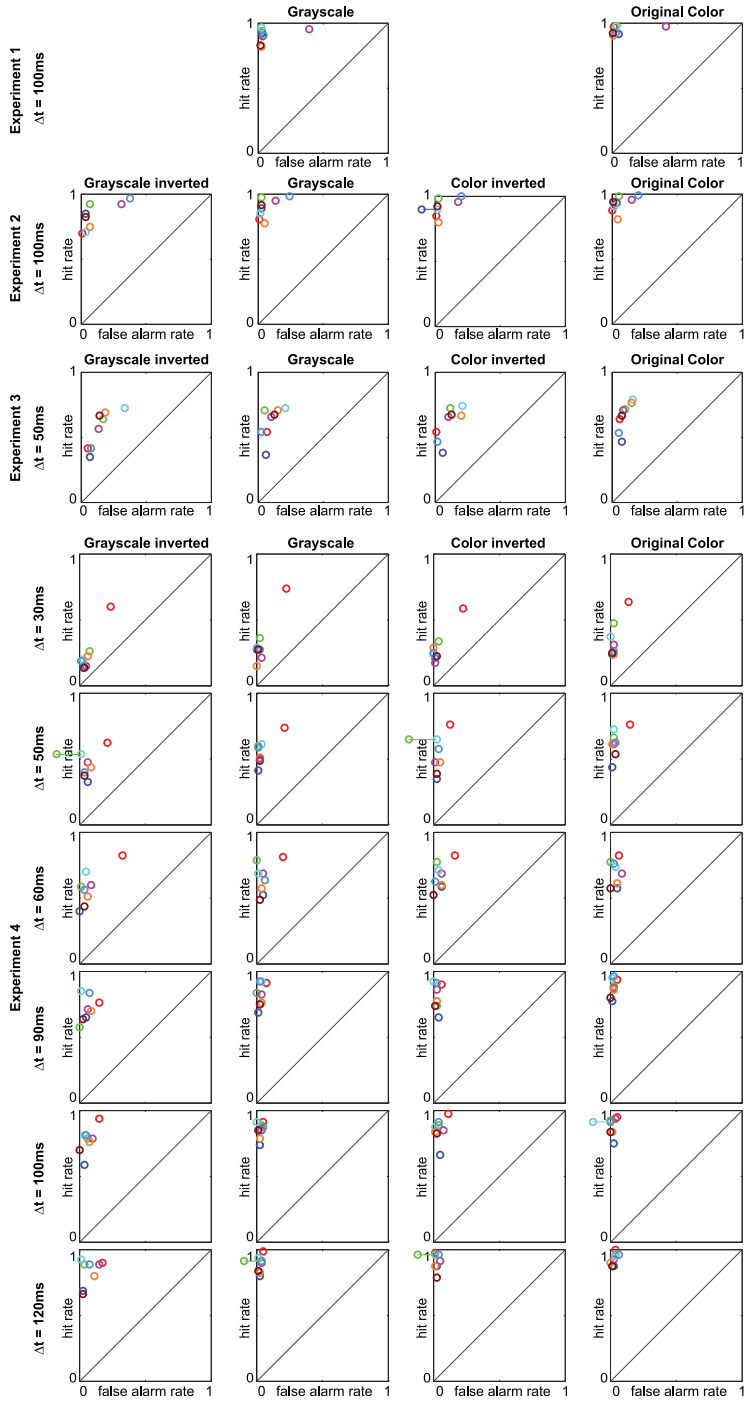

Figure 6. Raw data, hit rates, and false-alarm rates. All hit and false-alarm rates for all experiments, conditions, and SOAs. Within each experiment, the same color denotes the same observer. Responses 1 and 2 are counted as false alarms and hits for this representation. If data points were exactly overlapping, the covered data point was moved horizontally out of the axes and connected to its original location with a thin line.

demonstrated that lag-1 sparing occurred only when T1 and T2 were at the same location, a condition that may be violated in complex scenes where the target does not fill the full image. In addition, lag-1 sparing decreases for lower similarity between T1 and T2 (Visser, Davis, & Ohan, 2009). On a basic feature level, animal targets can be rather dissimilar, and the subordinate categorical similarity does not seem to be of relevance for lag-1 sparing in our experiment (detection at lag 1 was virtually identical, no matter if T1 and T2 were of the same or different categories). Finally, there are a number of other conditions under which lag-1 sparing is not observed, for example, when no short-term consolidation takes place (Dell' Acqua, Jolicœur, Pascali, & Pluchino, 2007) or when T1 is masked (Martin & Shapiro, 2008). So while lag-1 sparing is widely considered a hallmark of the AB, there are AB conditions in which no lag-1 sparing is observed, and therefore lag-1 sparing is not necessarily considered indicative of an AB effect (MacLean & Arnell, 2012).

In our paradigm, responses were unspeeded and we did not measure reaction times. Hence, we cannot fully rule out the possibility that for some targets, especially at short SOAs, color could have sped up the responses without affecting accuracy, as has been reported earlier (Oliva & Schyns, 2000; Rousselet et al., 2005). Since decreased reaction times are associated with increased confidence (Henmon, 1911), such speeding up could, however, also be related to increased subjective confidence for color as compared to grayscale sequences, which has been reported earlier even in the absence of a performance difference (Yao & Einhäuser, 2008). It should be noted, however, that our results of increased hit rates cannot be explained by a shift in criteria towards more liberal responses, since false-alarm rates were indistinguishable between conditions.

Whether detection and recognition are based on the same underlying process is a matter of debate. Grill-Spector and Kanwisher (2005) found the same performance and reaction times in detection and basic-level categorization tasks and concluded that figure–ground segmentation leading to detection and basic-level categorization are closely linked and mediated by one mechanism. In turn, this hypothesis has been challenged by subsequent studies that found better performance in detection than in categorization in basic-level categorization tasks (Bowers & Jones, 2008; Mack, Gauthier, Sadr, & Palmeri, 2008; Mack & Palmeri, 2010). It has been shown that both mechanisms can be selectively manipulated (Mack et al., 2008; Mack & Palmeri, 2010), and thus there is no intrinsic link between them. Here we find—consistent over all experiments and SOAs—that color has little influence on recognition, once the target has been detected. In contrast, the gray-inverted condition shows that luminance information influences recognition for detected targets. Since there are more false alarms for the gray-inverted than for any other condition, this result could still be explained by an increased number of guesses within the population of hits and thus decreased recognition performance. However, a similar effect is observed in all conditions of Experiment 4 when decreasing SOAs, which does not affect false-alarm rates. Decreasing SOAs not only reduces performance (in terms of d') but also further reduces the fraction of correctly recognized targets among the correctly detected ones. This argues against entirely overlapping mechanisms for detection and recognition. Nonetheless, even for short SOAs and the gray-inverted condition, recognition for detected targets is clearly above chance (>60%, with chance level at 25%). This offers an alternative explanation that, at least for high performance at larger SOAs, might contribute to the strong coupling between detection and recognition: For difficult targets (those only detected at large SOAs), the report of a detection may depend on some subordinate recognition. This view is supported by the conservative criterion nearly all observers apply for detection under difficult conditions (Figure 6). Whether distinct mechanisms or not, however, our data clearly show that color influences recognition and detection alike, such that once a target has been detected, the probability of it being correctly recognized does not depend on the presence of color.

Keywords: attentional blink, color, rapid serial visual processing, RSVP, natural scenes

Acknowledgments

This work was supported by the German Research Foundation (DFG; grants: EI 852/1, EI 852/3, IRTG-1901, and SFB/TRR135).

Commercial relationships: none.
Corresponding author: Svenja Marx.
Email: svenja.marx@physik.uni-marburg.de.
Address: FB Physik, AG Neurophysik, Philipps-University, Marburg, Germany.

References

Biederman, I. (1972, July 7). Perceiving real-world scenes. *Science, 177*(4043), 77–80.

Bowers, S. J., & Jones, K. W. (2008). Detecting objects is easier than categorizing them. *Quarterly Journal of Experimental Psychology, 61*(4), 552–557.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 109–127.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers, 34*(4), 613–617.

Dell' Acqua, R., Jolicœur, P., Pascali, A., & Pluchino, P. (2007). Short-term consolidation of individual identities leads to lag-1 sparing. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 593–609.

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: A study in monkeys and humans. *Vision Research, 40*(16), 2187–2200.

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2010). Key visual features for rapid categorization of animals in natural scenes. *Frontiers in Psychology, 1,* 21, doi:10.3389/fpsyg.2010.00021.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology, 357*(1), 241–265.

Einhäuser, W., Koch, C., & Makeig, S. (2007). The duration of the attentional blink in natural scenes depends on stimulus category. *Vision Research, 47*(5), 597–607.

Elder, J. H., & Velisavljević, L. (2009). Cue dynamics underlying rapid detection of animals in natural scenes. *Journal of Vision, 9*(7):7, 1–20, http://journalofvision.org/content/9/7/7, doi:10.1167/9.7.7. [PubMed] [Article]

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1476–1492.

Fabre-Thorpe, M., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *Neuroreport, 9*(2), 303–308.

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision, 7*(1):10, 1–29, http://journalofvision.org/content/7/1/10, doi:10.1167/7.1.10. [PubMed] [Article]

Fei-Fei, L., VanRullen, R., Koch, C., & Perona, P. (2005). Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. *Visual Cognition, 12*(6), 893–924.

Frey, H.-P., Honey, C., & König, P. (2008). What's color got to do with it? The influence of color on visual attention in different categories. *Journal of Vision, 8*(14):6, 1–17, http://journalofvision.org/content/8/14/6, doi:10.1167/8.14.6. [PubMed] [Article]

Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology, 10*(13), 805–808.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science, 16*(2), 152–160.

Hansen, T., & Gegenfurtner, K. R. (2009). Independence of color and luminance edges in natural scenes. *Visual Neuroscience, 26*(1), 35–49.

Henmon, V. A. C. (1911). The relation of the time of a judgment to its accuracy. *Psychological Review, 18*(3), 186–201.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46*(11), 1762–1776.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, USA, 99*(14), 9596–9601.

Mack, M. L., Gauthier, I., Sadr, J., & Palmeri, T. J. (2008). Object detection and basic-level categorization: Sometimes you know it is there before you know what it is. *Psychonomic Bulletin & Review, 15*(1), 28–35.

Mack, M. L., & Palmeri, T. J. (2010). Decoupling object detection and categorization. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1067–1079.

MacLean, M. H., & Arnell, K. M. (2012). A conceptual and methodological framework for measuring and modulating the attentional blink. *Attention, Perception & Psychophysics, 74*(6), 1080–1097.

Macmillan, N. A. (1993). Signal detection theory as data analysis method and psychological decision model. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: methodological issues* (pp. 21–57). Hillsdale, NJ: Lawrence Erlbaum Associates.

Martin, E. W., & Shapiro, K. L. (2008). Does failure to mask T1 cause lag-1 sparing in the attentional blink? *Perception & Psychophysics, 70*(3), 562–570.

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences, 29*(6), 317–322.

Meng, M., & Potter, M. C. (2008). Detecting and

remembering pictures with and without visual noise. *Journal of Vision, 8*(9):7, 1–10, http://www.journalofvision.org/content/8/9/7, doi:10.1167/8.9.7. [PubMed] [Article]

Motter, B. C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience, 14*(4), 2178–2189.

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41*(2), 176–210.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology, 81*(1), 10–15.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849–860.

Reddy, L., Wilken, P., & Koch, C. (2004). Face-gender discrimination is possible in the near-absence of attention. *Journal of Vision, 4*(2):4, 106–117, http://www.journalofvision.org/content/4/2/4, doi:10.1167/4.2.4. [PubMed] [Article]

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience, 5*(7), 629–630.

Rousselet, G. A., Joubert, O., & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes? *Visual Cognition, 12*(6), 852–877.

Skiera, G., Petersen, D., Skalej, M., & Fahle, M. (2000). Correlates of figure-ground segregation in fMRI. *Vision Research, 40*(15), 2047–2056.

Spence, I., Wong, P., Rusan, M., & Rastegar, N. (2006). How color enhances visual memory for natural scenes. *Psychological Science, 17*(1), 1–6.

Thorpe, S. J., Delorme, A., & VanRullen, R. (2001). Spike-based strategies for rapid processing. *Neural Networks, 14*(6), 715–725.

Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520–522.

Thorpe, S. J., & Gautrais, J. (1997). Rapid visual processing using spike asynchrony. In M. C. Mozer, M. Jordan, & T. Petsche (Eds.), *Advances in neural information processing systems 9* (pp. 901–907). Cambridge, MA: MIT Press.

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception, 30*(6), 655–668.

Visser, T. A. W., Davis, C., & Ohan, J. L. (2009). When similarity leads to sparing: Probing mechanisms underlying the attentional blink. *Psychological Research, 73,* 327–335.

Visser, T. A. W., Zuvic, S. M., Bischof, W., & Di Lollo, V. (1999). The attentional blink with targets in different spatial locations. *Psychonomic Bulletin & Review, 6*(3), 432–436.

Wichmann, F. A., Braun, D. I., & Gegenfurtner, K. R. (2006). Phase noise and the classification of natural images. *Vision Research, 46*(8), 1520–1529.

Wichmann, F. A., Drewes, J., Rosas, P., & Gegenfurtner, K. R. (2010). Animal detection in natural scenes: critical features revisited. *Journal of Vision, 10*(4):6, 1–27, http://www.journalofvision.org/content/10/4/6, doi:10.1167/10.4.6. [PubMed] [Article]

Wichmann, F. A., Sharpe, L. T., & Gegenfurtner, K. R. (2002). The contributions of color to recognition memory for natural scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 509–520.

Wurm, L. H., Legge, G. E., Isenberg, L. M., & Luebker, A. (1993). Color improves object recognition in normal and low vision. *Journal of Experimental Psychology: Human Perception and Performance, 19*(4), 899–911.

Yao, A. Y. J., & Einhäuser, W. (2008). Color aids late but not early stages of rapid natural scene recognition. *Journal of Vision, 8*(16):12, 1–13, http://www.journalofvision.org/content/8/16/12, doi:10.1167/8.16.12. [PubMed] [Article]

Zhang, W., & Luck, S. J. (2009). Feature-based attention modulates feedforward visual processing. *Nature Neuroscience, 12*(1), 24–25.

Appendix A: Individual responses

With a few exceptions, observers in all experiments show consistent patterns with respect to their performance. In general, all are conservative (making more misses than false alarms), and observers with comparably liberal criteria tend to remain so across all conditions (Figure 6). Considering all nine combinations of ground truth and response for Experiments 1–3, the incidence of double false alarms is small (Figure A1) and—except for a few individuals and conditions—so are false alarms in single-target trials (i.e., two reported targets where one is correct).

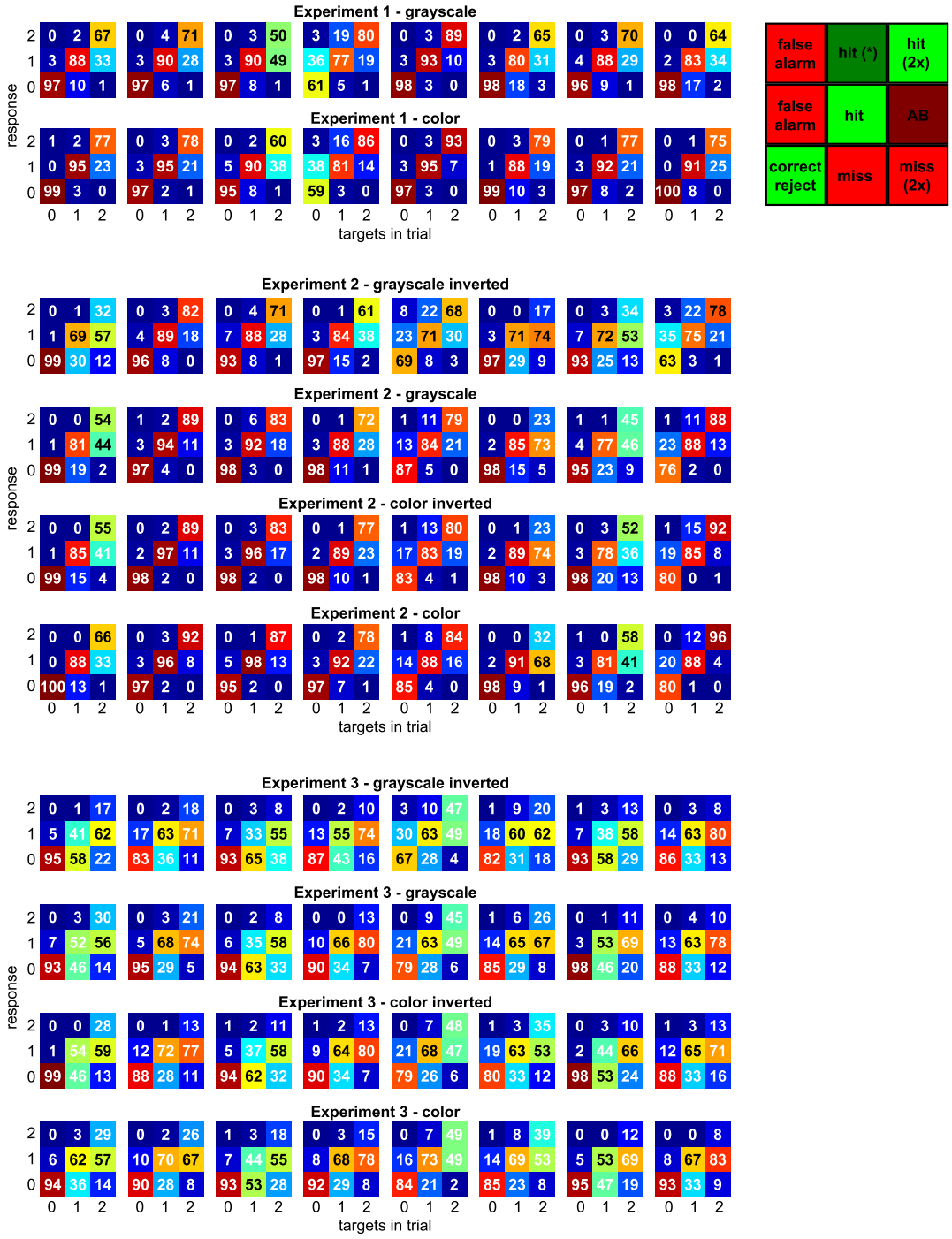

Figure A1. Raw data, all combinations of ground truth and response. For Experiments 1–3, there were nine combinations of correct responses (truth) and actual responses (response). For each individual, experiment, and condition, the raw percentage of responses for the respective truth are color-coded and provided (i.e., columns sum to 100%). The large matrix on the top right defines signal detection theory (SDT) terms. Note that "hit(*)" contains 1 hit and 1 false alarm (truth 1, response 2) and yields the two different definitions of hit used in the article.

Journal of Vision (2014) 14(14):4, 1–18

Appendix B: Analysis of hits and false alarms

In all experiments, we find main effects of condition for hits (in either definition) and for false alarms, while in Experiment 4 we additionally find a main effect of SOA for hits though not for false alarms (Table 1) .There are significantly more hits (in either definition) for color than for grayscale images across all experiments and SOAs, with the exception of the shortest SOA (30 ms) in Experiment 4 (Table 3). Interestingly, such a difference cannot be identified for false alarms (Table 4). With respect to hit rates, the color-inverted condition is different from grayscale and indistinguishable from original color at the 100-ms SOA of Experiment 2 when hits are defined as reporting at least one target in one-target trials (Table 3). If we instead restrict hits to correct responses (response = 1 for one-target streams), the picture reverses and now the color-inverted condition is indistinguishable from grayscale but yields significantly fewer hits than the original-color condition. For the 50 ms of Experiment 3, this reversed pattern holds for either definition. This underlines the importance of conducting Experiment 4, where zero targets and one were the only possible response options. Considering hits alone, the color-inverted condition is indistinguishable from gray for all SOAs of Experiment 4 and different from color for 30, 50, and 90 ms (Table 3). In the gray-inverted condition, significantly fewer hits than in any other condition in all experiments can be observed: response > 0: all $t(7) > 2.74$, all $ps < 0.03$; for response $= 1$: all $t(7) > 2.81$, all $ps < 0.026$.

Appendix C: Detailed analysis of the attentional blink

In the separate analysis of color and grayscale sequences, we find a significant main effect of lag on the probability that both targets are detected: gray: $F(4, 28) = 18.63$, $p = 1.4 \times 10^{-7}$; color: $F(4, 28) = 33.26$, $p = 2.9 \times 10^{-10}$). There is a monotonic increase in performance up to lag 4 (Figure 4A). Pairwise post hoc tests show in both color conditions that lags 1, 2, and 3 are significantly different from lag 7, all $ts > 2.5$, all $ps < 0.05$, while lag 4 is not different from lag 7: gray: $t(7) = 0.48$, $p = 0.64$; color: $t(7) = 1.05$, $p = 0.33$. Using the squared single-target hit rate of the respective color category as a baseline, lags 1, 2, and 3 are different from this baseline, all $ts > 2.5$, all $ps < 0.05$, while lags 4 and 7 are indistinguishable from the baseline, all $ts < 2.01$, all $ps > 0.08$. Similarly, in Experiments 2 and 3, where we tested only lag 2, there is a significant difference between the two-target hit rate and the baseline in each color condition: Experiment 2: all $ts > 2.4$, all $ps < 0.04$, Figure 4B; Experiment 3: all $ts > 3.24$, all $ps < 0.014$, Figure 4C.

Appendix D: Single-target recognition given detection in Experiments 2, 3, and 4

In Experiments 2 and 3, the main effect of color condition on recognition-given-detection performance is solely explained by the gray-inverted category: While the gray-inverted condition is different from all other conditions—Experiment 2: $ts > 6.06$, $ps < 5.1 \times 10^{-4}$; Experiment 3: $ts > 4.22$, $ps < 0.0039$—there are no pairwise differences between any of the other conditions: Experiment 2: all $ts < 1.71$, $ps > 0.13$; Experiment 3: $ts < 1.66$, $ps > 0.14$. In Experiment 4, the main effects of SOA and color on this measure also result almost entirely from the difference between the gray-inverted and all other conditions, all $ts > 2.99$, all $ps < 0.02$, with the exception of the difference between gray inverted and color at an SOA of 100 ms, which is not significant, $t(7) = 1.54$, $p = 0.17$, and the difference between grayscale and color inverted at an SOA of 120 ms, which is significant, $t(7) = 2.90$, $p = 0.023$.

Appendix E: Recognition in two-target trials with one hit

When one target is detected and one is missed, neither the probability that the reported target category matches the first target nor that it matches the second target depends on lag or color condition: T1, color condition: $F(1, 24) = 0.53$, $p = 0.49$; T1, lag: $F(4, 24) = 1.81$, $p = 0.16$; T1, interaction: $F(4, 24) = 0.61$, $p = 0.66$; T2: $F(1, 24) = 0.29$, $p = 0.61$; $F(4, 24) = 1.61$, $p = 0.20$; $F(4, 24) = 0.50$, $p = 0.73$—observer #5 could not be included in this particular analysis, as she had no miss trial in one of the two-target conditions (lag 4, color). In Experiments 2 and 3 similarly, given that one target is detected in a two-target stream, the probability that it matched T1 or T2 does not depend on color condition: Experiment 2, T1: $F(3, 21) = 0.45$, $p = 0.72$; T2: $F(3, 21) = 2.65$, $p = 0.075$; Experiment 3, T1: $F(3, 21) = 1.56$, $p = 0.23$; T2: $F(3, 21) = 0.81$, $p = 0.50$. Hence color has no effect on attention or recognition beyond the effect that is already explained by detection.

	SOA	Gray inverted	Grayscale	Color inverted	Original color	Gray vs. original color	Gray vs. color inverted	Color inverted vs. original color
Response > 0								
Experiment 1	100 ms		90.6% ± 5.5%		94.6% ± 3.1%	$\mathbf{t(7) = 3.01}$ $\mathbf{p = 0.020}$		
Experiment 2	100 ms	83.1% ± 10.4%	89.4% ± 7.7%	91.5% ± 6.7%	92.5% ± 6.1%	$\mathbf{t(7) = 3.84}$ $\mathbf{p = 0.0064}$	$\mathbf{t(7) = 3.54}$ $\mathbf{p = 0.0095}$	$t(7) = 2.11$ $p = 0.072$
Experiment 3	50 ms	55.9% ± 14.5%	61.6% ± 12.4%	60.7% ± 12.9%	66.2% ± 11.2%	$\mathbf{t(7) = 2.93}$ $\mathbf{p = 0.022}$	$t(7) = 0.71$ $p = 0.50$	$\mathbf{t(7) = 3.55}$ $\mathbf{p = 0.0094}$
Response $= 1$								
Experiment 1	100 ms		86.1% ± 5.8%		90.9% ± 4.7%	$\mathbf{t(7) = 4.47}$ $\mathbf{p = 0.0029}$		
Experiment 2	100 ms	76.5% ± 7.26%	85.7% ± 5.2%	87.1% ± 5.8%	89.7% ± 5.1%	$\mathbf{t(7) = 5.16}$ $\mathbf{p = 0.0013}$	$t(7) = 1.45$ $p = 0.19$	$\mathbf{t(7) = 6.52}$ $\mathbf{p = 3 \times 10^{-4}}$
Experiment 3	50 ms	51.9% ± 12.6%	58.1% ± 11.0%	58.3% ± 12.3%	63.2% ± 9.8%	$\mathbf{t(7) = 3.81}$ $\mathbf{p = 0.0066}$	$t(7) = 0.13$ $p = 0.90$	$\mathbf{t(7) = 3.93}$ $\mathbf{p = 0.0056}$
Experiment 4								
	30 ms	23.8% ± 15.3%	32.2% ± 17.9%	28.9% ± 13.0%	34.8% ± 14.3%	$t(7) = 0.94$ $p = 0.38$	$t(7) = 1.17$ $p = 0.28$	$\mathbf{t(7) = 2.38}$ $\mathbf{p = 0.049}$
	50 ms	46.4% ± 9.9%	55.6% ± 9.9%	54.1% ± 14.3%	62.2% ± 10.2%	$\mathbf{t(7) = 4.93}$ $\mathbf{p = 0.0017}$	$t(7) = 0.86$ $p = 0.42$	$\mathbf{t(7) = 3.97}$ $\mathbf{p = 0.0054}$
	60 ms	57.8% ± 13.8%	65.0% ± 11.7%	66.7% ± 10.1%	69.4% ± 9.7%	$\mathbf{t(7) = 2.70}$ $\mathbf{p = 0.03}$	$t(7) = 1.88$ $p = 0.10$	$t(7) = 1.54$ $p = 0.17$
	90 ms	71.6% ± 9.7%	82.8% ± 9.0%	81.3% ± 10.1%	88.1% ± 6.9%	$\mathbf{t(7) = 5.49}$ $\mathbf{p = 0.00091}$	$t(7) = 1.09$ $p = 0.31$	$\mathbf{t(7) = 4.00}$ $\mathbf{p = 0.0052}$
	100 ms	77.3% ± 10.0%	85.2% ± 6.1%	85.3% ± 9.1%	88.3% ± 6.8%	$\mathbf{t(7) = 2.82}$ $\mathbf{p = 0.026}$	$t(7) = 0.094$ $p = 0.93$	$t(7) = 2.04$ $p = 0.081$
	120 ms	83.0% ± 10.2%	88.9% ± 6.3%	91.4% ± 6.5%	93.1% ± 4.5%	$\mathbf{t(7) = 4.78}$ $\mathbf{p = 0.0020}$	$t(7) = 1.74$ $p = 0.13$	$t(7) = 1.55$ $p = 0.16$

Table 3. Hit rates and post hoc comparisons of interest. Bold type indicates a significant effect; all percentages are mean ± standard deviation.

2.3 Rapid serial processing of natural scenes: Color modulates detection but neither recognition nor the attentional blink

	SOA	Gray inverted	Grayscale	Color inverted	Original color	Grayscale vs. original color	Grayscale vs. color inverted	Color inverted vs. original color
Experiment 1	100 ms		7.4% ± 12.9		7.2% ± 13.8%	$t(7) = 0.27$ $p = 0.80$		
Experiment 2	100 ms	11.6% ± 14.2%	6.6% ± 8.2%	6.0% ± 7.9%	6.3% ± 7.3%	$t(7) = 0.42$ $p = 0.68$	$t(7) = 0.61$ $p = 0.56$	$t(7) = 0.36$ $p = 0.73$
Experiment 3	50 ms	14.4% ± 9.1%	9.8% ± 6.1%	10.4% ± 7.5%	9.4% ± 4.0%	$t(7) = 0.34$ $p = 0.75$	$t(7) = 0.47$ $p = 0.66$	$t(7) = 0.74$ $p = 0.48$
Experiment 4	30 ms	6.6% ± 7.3%	4.0% ± 7.6%	3.9% ± 7.6%	3.3% ± 4.3%	$t(7) = 0.65$ $p = 0.54$	$t(7) = 0.28$ $p = 0.78$	$t(7) = 0.49$ $p = 0.64$
	50 ms	6.6% ± 6.5%	4.5% ± 6.8%	4.1% ± 3.6%	4.1% ± 4.5%	$t(7) = 0.50$ $p = 0.63$	$t(7) = 0.36$ $p = 0.73$	$t(7) = 0$ $p = 1$
	60 ms	7.7% ± 10.4%	5.5% ± 6.2%	5.3% ± 5.0%	3.9% ± 3.0%	$t(7) = 0.80$ $p = 0.45$	$t(7) = 0.14$ $p = 0.89$	$t(7) = 1.05$ $p = 0.33$
	90 ms	5.8% ± 4.8%	3.1% ± 2.2%	2.7% ± 1.8%	2.2% ± 1.5%	$t(7) = 1.53$ $p = 0.17$	$t(7) = 0.63$ $p = 0.55$	$t(7) = 1.16$ $p = 0.28$
	100 ms	6.4% ± 4.6%	3.0% ± 1.8%	4.5% ± 3.4%	1.6% ± 2.0%	$t(7) = 2.05$ $p = 0.080$	$t(7) = 1.72$ $p = 0.13$	**$t(7) = 4.46$** **$p = 0.0029$**
	120 ms	7.7% ± 6.3%	3.0% ± 1.3%	2.3% ± 1.4%	3.0% ± 1.9%	$t(7) = 0$ $p = 1$	$t(7) = 1.72$ $p = 0.13$	$t(7) = 0.83$ $p = 0.43$

Table 4. False-alarm rates and post hoc comparisons of interest. Bold type indicates a significant effect; all percentages are mean ± standard deviation.

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 10 February 2015
doi: 10.3389/fnhum.2015.00046

Effects of aging on eye movements in the real world

Stefan Dowiasch *, Svenja Marx, Wolfgang Einhäuser and Frank Bremmer

Department of Neurophysics, Philipps-University Marburg, Marburg, Germany

Edited by:
John J. Foxe, Albert Einstein College of Medicine, USA

Reviewed by:
Miguel Escudero, University of Seville, Spain
Sheree McCormick, Manchester Metropolitan University, UK
Nicolas Carvalho, University of Franche-Comté, France

****Correspondence:***
Stefan Dowiasch, Department of Neurophysics, Philipps-University Marburg, Karl-von-Frisch-Straße 8a, 35043 Marburg, Germany
e-mail: stefan.dowiasch@physik.uni-marburg.de

The effects of aging on eye movements are well studied in the laboratory. Increased saccade latencies or decreased smooth-pursuit gain are well established findings. The question remains whether these findings are influenced by the rather untypical environment of a laboratory; that is, whether or not they transfer to the real world. We measured 34 healthy participants between the age of 25 and 85 during two everyday tasks in the real world: (I) walking down a hallway with free gaze, (II) visual tracking of an earth-fixed object while walking straight-ahead. Eye movements were recorded with a mobile light-weight eye tracker, the EyeSeeCam (ESC). We find that age significantly influences saccade parameters. With increasing age, saccade frequency, amplitude, peak velocity, and mean velocity are reduced and the velocity/amplitude distribution as well as the velocity profile become less skewed. In contrast to laboratory results on smooth pursuit, we did not find a significant effect of age on tracking eye-movements in the real world. Taken together, age-related eye-movement changes as measured in the laboratory only partly resemble those in the real world. It is well-conceivable that in the real world additional sensory cues, such as head-movement or vestibular signals, may partially compensate for age-related effects, which, according to this view, would be specific to early motion processing. In any case, our results highlight the importance of validity for natural situations when studying the impact of aging on real-life performance.

Keywords: eye movements, aging, real-world gaze, natural environment, self-motion, saccades, tracking eye-movements

INTRODUCTION

As we are getting older, the function of the visual system appears to deteriorate. Not only does visual acuity decline in the elderly, but perception (e.g., Billino et al., 2008; Lich and Bremmer, 2014) and eye-movement parameters are also altered (Morgan, 1993). Increased saccadic latencies (Abel et al., 1983; Moschner and Baloh, 1994; Munoz et al., 1998; Klein et al., 2000) and decreased smooth-pursuit gain (Moschner and Baloh, 1994; Ross et al., 1999) are common findings in the literature, while the results for other oculomotor parameters like saccade peak-velocity are inconclusive. Some studies found a decrease during senescence (Warabi et al., 1984; Sharpe and Zackon, 1987; Irving et al., 2006), whereas others could not show a significant correlation of age and saccade peak velocity (Henriksson et al., 1980; Munoz et al., 1998).

In the last decades, the study of eye movements has increased in relevance as gaze serves as an easily accessible, reliable, safe and fast proxy for cognitive processes and as tool to identify possible functional impairments of the brain (Leigh and Zee, 2006). As an example, the measurement of saccade amplitude and velocity offers an indication of the functionality of the saccade generating circuitry in the brainstem (Sparks, 2002). Certain eye-movement characteristics may extend the knowledge of the mechanism underlying some neurological and psychiatric diseases (Gooding and Basso, 2008; Pinkhardt et al., 2008; Marx et al., 2012; Dowiasch et al., 2014), and might in the long-run, support diagnosis in the clinical routine.

Self-motion through an environment induces one of the most fundamental causes for differences between eye movements in the laboratory and the real world. For example, during walking, the eye-movement system encounters distinct demands as compared to sitting still in the laboratory, which is reflected in qualitatively different oculomotor behavior ('t Hart et al., 2009; 't Hart and Einhäuser, 2012). For example, keeping the eyes on a target that is stationary in the world turns from a mere fixation in the laboratory into a tracking eye-movement during self-motion (Niemann et al., 1999), since the projection of every location in our visual field moves across the retina. Likewise, smooth-pursuit eye-movements as performed in the laboratory are often accompanied by head movements and vestibular-ocular reflexes during free real-world movement. Therefore, these eye movements have to integrate self-motion information in order to operate optimally. At the cortical level, this leads to a massive involvement of areas of the dorsal pathway where the processing of self-motion signals primarily takes place (Bremmer et al., 2000). Especially areas like the ventral intraparietal area (VIP; Bremmer et al., 2001, 2002a; Britten, 2008; Wall and Smith, 2008; Chen et al., 2011) and the medial superior temporal area (MST; Duffy and Wurtz, 1991; Bremmer et al., 1999; Gu et al., 2008; Pitzalis et al., 2013) get activated not only by visual but also by vestibular self-motion signals.

Despite the increasing interest in real-word eye-tracking and despite the abundance of literature on how eye movements in the laboratory are affected in healthy aging, to the best of our knowledge, no study has addressed the effects of healthy aging on real-world eye-movement behavior. Such a transfer to the real world, however, seems particularly important, as an increasing number of studies on eye movements in real-world environments and during everyday tasks (e.g., Land et al., 1999; Hayhoe and Ballard, 2005) raise substantial doubt as to whether results from the laboratory can be directly transferred to real-world scenarios ('t Hart et al., 2009; Foulsham et al., 2011; Tatler et al., 2011; 't Hart and Einhäuser, 2012). Since these studies have typically been performed with small samples and within a homogenous age group, they left possibly existing effects of age on real-world eye-movement behavior so far unaddressed. Patient studies on oculomotor deficiencies in disease, in turn, typically include a set of age-matched healthy controls and often span a wider age range, but do not typically assess the factor age explicitly. In this study, we draw on such control data from earlier patient studies to close the gap and test if and if so to what extent age relate to eye movements in a comparably unconstrained real-world setting. Specifically, we tracked participants' eye movements while they walked in a corridor either looking around freely or tracking a stationary target on the floor during walking. Since re-inviting the same cohort of participants to laboratory measurements was not feasible, we compared the real-world data in our study to common findings from laboratory studies reported in the literature. It is self-evident that there is no 1-to-1 mapping between such tasks. These limitations notwithstanding, any differences between our results and studies performed in the laboratory may suggest how age-related changes in the healthy brain affect gaze behavior in real-life situations. Such findings would underline the need for addressing real-world tasks to complement laboratory measurements towards a full understanding of the mechanisms underlying oculomotor changes during healthy aging and might point towards new research objectives of future studies.

FIGURE 1 | Illustration of a typical scene during calibration and the two tasks. Images were taken from the head mounted camera of the ESC. The red square indicates the current gaze position of a participant. **(A)** Calibration: Fixating stationary targets with a fixed distance of 7° as projected by a head-fixed laserpointer of the ESC (enhanced in this figure for visualization). **(B)** Task I: Walking down the hallway with free gaze and **(C)** Task II: visually tracking two stationary targets on the floor while walking straight-ahead.

METHODS

SUBJECTS

The eye movements of 34 participants (31 male, 3 female) between the age of 25 and 85 (mean = 46y ± 18.5y) were analyzed during two everyday tasks in the real world. All participants were originally recruited as healthy controls for patient studies on eye movements in natural environments (Marx et al., 2012; Dowiasch et al., 2014). Each participant had normal or corrected to normal vision and no history of neurological or psychiatric disease. Two of the tasks in these two studies were identical and are used for the present analysis. Both studies were approved by the local ethics committee and were in accordance with the Declaration of Helsinki. All participants gave their written informed consent.

DATA ACQUISITION

Binocular eye-in-head movements were recorded with a mobile light-weight eye tracker, the EyeSeeCam (ESC), at a sampling rate of 280 Hz, a spatial resolution of 0.02° and a precision of about 0.1° (Schneider et al., 2009). This allowed us to record and analyze saccadic eye-movements during walking with free gaze reliably. The ESC records a head-centered video with a head-fixed camera and provides a video sequence obtained from a movable camera (GazeCam) which follows the gaze of the participant with a constant latency of about 10 ms (Schneider et al., 2009). Before each measurement the system

was calibrated by matching the gaze direction of the subject with the position of 5 predefined targets, which were projected to a plain wall at a distance of 2 m by a head fixed laser pointer (**Figure 1A**). The mean error threshold for calibration was set to 0.5°. After successful calibration, the participants were asked to perform two different tasks in an indoor environment and to act as they normally would throughout these tasks.

BEHAVIORAL TASKS

In the first task participants had to walk along a hallway for about 35 m with free gaze and no additional instruction (**Figure 1B**). For the second task, they were asked to visually track a stationary orange colored spot with a diameter of 10 cm on the ground (dark green carpet) while walking towards it, starting about 10 m away from the spot (**Figure 1C**). Each participant was free to choose his/her own walking speed. The duration of the full set of measurements including setup (∼1 min), accustoming (at least 2 min; participants indicated when they feel ready to perform the tasks) and calibration (∼1 min) of the eye tracker ranged from 5 to 10 min per participant.

Recorded eye-position data and video sequences were analyzed offline using MATLAB 2010b (The MathWorks, Inc., Natick, Massachusetts, USA). In a first step, raw eye-position data was inspected for blinks or other recording artifacts due e.g., to reflections by external light sources. Blinks were classified by the absence of more than 5 samples (18 ms) and eye traces were cleaned for blink artifacts by deleting 8 samples (29 ms) before the start of a blink and 12 samples (43 ms) after a blink. Saccades were detected if eye velocity was higher than 100°/s for at least 3 consecutive samples and if the eyes moved more than 0.5° in this time period. This conservative threshold guaranteed a low false-positive rate for saccade detection, since eye movements during real-life measurements contain extensive dynamics (e.g., due to vestibulo-ocular reflexes or vergence eye movements) and are generally noisier than under controlled laboratory settings. Furthermore, a main-sequence analysis (peak velocity/amplitude) of thus defined saccades was performed by computing the power function fit ($v_{\text{peak}} = K * amplitude^{L}$) and its corresponding 95% confidence interval for each subject (Bahill et al., 1975). All saccades outside this interval were classified as outliers and were not considered for further analysis. Additionally, the saccade velocity profile was characterized by using the q-value, which is defined as the ratio of peak- and mean velocity ($v_{\text{peak}}/v_{\text{mean}}$; Inchingolo et al., 1987). Finally, saccade amplitude, mean- and peak-velocities were separately analyzed for each of the four cardinal directions (right, left, up, and down). Therefore only saccades with a mean velocity component of more than 100°/s in one of the four directions were considered for analysis to exclude saccades with no specific cardinal direction.

The tracking performance of each participant was quantified by eye-in-head gain values (eye velocity divided by target velocity) and the RMSE (root mean square error) of the retinal target velocity. The rationale for choosing the RMSE, just as for the gain, was its wide use as a global measure of pursuit performance (Smyrnis, 2008) and its good test-retest reliability (Gooding et al., 1994). As a first step in analyzing tracking, all tracking segments were cleaned from saccadic artifacts such as catch-up saccades to analyze the smooth tracking phase only. Since subjects were free to move their eyes in Task I, they typically tracked multiple objects during their way through the hallway ("spontaneous tracking"). In this task only tracking segments longer than 200 ms were considered for further analysis. Accordingly, the reference velocity (target velocity) had to be determined individually for each subject and each eye-movement trajectory. To do so, we computed the optical flow field (Gautama and Van Hulle, 2002) from the head centered video recorded by the ESC. Target velocity was considered the velocity of the image part relative to the head which was tracked by the subjects' gaze. Due to technical issues in the recording of the head-centered video, (e.g., caused by blurred video or by considerable frame drops), the optic flow field of the recordings from 4 participants (age: 33, 50, 64, and 69) could not be computed reliably. These 4 subjects were excluded from optic-flow analysis and no free-viewing gain was computed.

In Task II the presence of specific tracking targets allowed us to determine target velocity as the temporal derivative of the target position in the head-centered scene. In addition, the GazeCam videos of the ESC could be used to calculate retinal target velocity as the temporal derivative of the target position within this retinocentric framework. The sum of all deviations from the optimal retinal target velocity (0°/s) corresponds to the RMSE. In this task, six participants (ages: 30, 33, 33, 50, 64, 74) did not show a sufficient tracking of the specified target (e.g., they ignored the target at all) and their data were therefore excluded from the evaluation of tracking performance (i.e., tracking-gain as well as RMSE). In addition the RMSE could not be evaluated precisely due to considerable frame drops in the GazeCam-videos of three other participants (ages: 30,32,53).

STATISTICS

The analyzed eye-movement parameters (saccade amplitude, saccade peak velocity, saccade mean velocity, and the q-value of the saccade velocity distribution in task I as well as tracking gain and RSME in task II) cannot be expected to follow a normal distribution (Land et al., 1999). Hence we used the non-parametric Mann-Whitney-U-Test (Mann and Whitney, 1947) for all statistical analyses. An alpha-level of 0.05 was used as threshold for significance. We characterized each subject's respective eye-movement parameter by the median (over saccades and tracking epochs, respectively) rather than by the mean. To calculate significance we performed a median-split analysis, comparing the older half of participants to the younger half. This resulted in a younger group ($n = 17$) with a mean age of 30.1 ± 3.0y and an older group ($n = 17$) with a mean age of 61.8 ± 12.8y. There was no statistical difference in mean age of the younger or older group when comparing the tasks in which participants were excluded and the full group of participants (freeviewing-gain: younger group ($n = 15$): mean age 29.5 ± 2.6y; older group ($n = 15$): mean age 60.2 ± 14.7y; tracking-gain: younger group ($n = 14$): mean age

Table 1 | Basic eye-movement parameters by task.

	Eye-movement parameter		U-test				
			Young (SD)	Old (SD)	*Z*-value	*p*-value	Effect-Size (AUC)
Task I (walking with free gaze)	median saccade amplitude [°]		4.14 (1.31)	3.25 (1.07)	1.894	0.058	0.692 [0.513 0.871]
	median saccade peak-velocity [°/s]		285.6 (51.4)	220.2 (24.2)	3.789	0.0002	0.882 [0.764 1.000]
	median saccade mean-velocity [°/s]		199.2 (29.1)	168.3 (17.8)	3.134	0.0017	0.817 [0.671 0.962]
	median q-value		1.40 (0.06)	1.31 (0.06)	3.479	0.0005	0.851 [0.719 0.984]
	main-sequence fit K-value		185.0 (24.7)	154.4 (22.8)	3.169	0.0015	0.820 [0.676 0.965]
	main-sequence fit L-value		0.335 (0.04)	0.359 (0.08)	−0.689	0.491	0.429 [0.235 0.624]
	mean saccade frequency [1/s]		3.03 (1.37)	1.83 (1.04)	2.582	0.010	0.761 [0.598 0.924]
	mean freeviewing-gain		1.31 (0.31)	1.56 (0.29)	−2.224	0.026	0.259 [0.077 0.441]
	mean blink rate [1/s]		0.474 (0.32)	0.801 (0.57)	−1.826	0.068	0.315 [0.135 0.495]
	median saccade peak-velocity [°/s]	1°–2° (22.5%)	202.9 (35.1)	187.4 (37.1)	2.032	0.042	0.706 [0.530 0.882]
	for certain saccade amplitudes	2°–3° (11.5%)	244.3 (52.9)	207.9 (28.1)	2.342	0.019	0.737 [0.568 0.906]
		3°–5° (14.7%)	251.6 (28.8)	226.9 (18.6)	2.962	0.003	0.799 [0.648 0.951]
		5°–8° (14.6%)	317.0 (34.8)	283.0 (24.9)	2.997	0.003	0.803 [0.652 0.953]
		8°–15° (15.3%)	413.6 (37.0)	363.9 (41.5)	2.514	0.012	0.855 [0.722 0.988]
Task II (tracking targets)	mean tracking-gain		0.957 (0.13)	0.865 (0.22)	1.264	0.206	0.643 [0.436 0.850]
	tracking RMSE		16.99 (6.97)	17.80 (4.90)	−0.680	0.497	0.417 [0.190 0.643]

Computed eye-movement parameters of the two tasks (Task I and II) and their corresponding statistics. The percentage behind each saccade amplitude window reflects the prevalence of a saccade of that particular amplitude. The distinct analysis of amplitude ranges of more than 15° could not be evaluated reliably, because some participants did not perform sufficient saccades within this domain.

29.7 ± 3.0y; older group (n = 14): mean age 61.6 ± 13.3y; RMSE: younger group (n = 12): mean age 29.5 ± 3.2y; older group (n = 13): mean age 62.2 ± 13.6y; all $p > 0.8$; U-Test). Additionally the effect size of each result was computed using the "Area under the receiver operating characteristic curve" (AUC or A') (Bamber, 1975). Similar to d', AUC can be understood as a measure of overlap of two distributions, with separability being minimal at a value of 0.5 and maximal at 0.0 or 1, respectively (Hentschke and Stüttgen, 2011). The 95% confidence interval for each effect size was calculated analytically (Hanley and McNeil, 1982). Finally, we report Pearson's linear correlation coefficient with the age of the participants for all eye-movement parameters and its corresponding significance level.

RESULTS

Thirty-four participants between the age of 25 to 85 performed two different oculomotor tasks in an indoor environment: (I) walking through a hallway with free gaze and (II) visually tracking a stationary object on the ground while walking straight ahead.

TASK I—WALKING WITH FREE GAZE

Saccade and blink rate

When walking along a hallway with no specific instructions we neither observed a significant difference in walking time between the younger (mean: 26.5 ± 2.7 s) and the older participants (mean: 24.1 ± 4.8 s; Z = 1.54; p = 0.124; U-test), nor a correlation of walking time and age ($r_{(32)} = -0.260$; p = 0.138). Yet, several eye-movement parameters depended on age. The frequency of saccades showed a significantly decrease in the older participants as compared to the younger (**Table 1**) and correlated with age ($r_{(32)} = -0.484$; p = 0.004; **Figure 2A**). There was a trend towards an increase in the frequency of eye blinks in older participants, but this trend did not reach statistical significance neither in the median-split analysis (**Table 1**), nor in the correlation analysis ($r_{(32)} = 0.313$; p = 0.07; **Figure 2B**).

Parameters of individual saccades (amplitude and velocities)

With respect to the parameters of individual saccades (amplitude and velocity measures), there was a trend towards smaller saccade amplitudes in the older participants, which however, did not reach significance (**Table 1**). Yet, we found a negative correlation between age and median saccade amplitude ($r_{(32)} = -0.396$; p = 0.02; **Figure 2C**). When analyzing the amplitudes of horizontal and vertical saccades separately only downward and leftward saccades showed a significant decrease in the older group (**Table 2**). Just as median saccade amplitude, saccade peak and mean velocity were negatively correlated with age (peak-velocity: $r_{(32)} = -0.626$; $p < 0.001$; **Figure 2D**; mean-velocity: $r_{(32)} = -0.567$; $p < 0.001$; **Figure 2E**). For those two parameters the median-split analysis showed a clearly significant decrease in the older population (**Table 1**). This was especially true for saccades in the horizontal direction (**Table 2**) and also trended to be significantly lower in the older group for downward saccades. Yet, upward saccades did not show any statistical differences between groups for either saccade mean- or peak-velocity (**Table 2**). Finally, the saccade distribution for older participants was less skewed, as reflected by the significantly lower median q-value (**Table 1**) and the negative correlation of age with q-value $r_{(32)} = -0.583$; $p < 0.001$; **Figure 2F**). When analyzing the standard deviations of the saccade parameters (amplitude, mean- and peak-velocities and q-value) in relation to the age of the participants, there

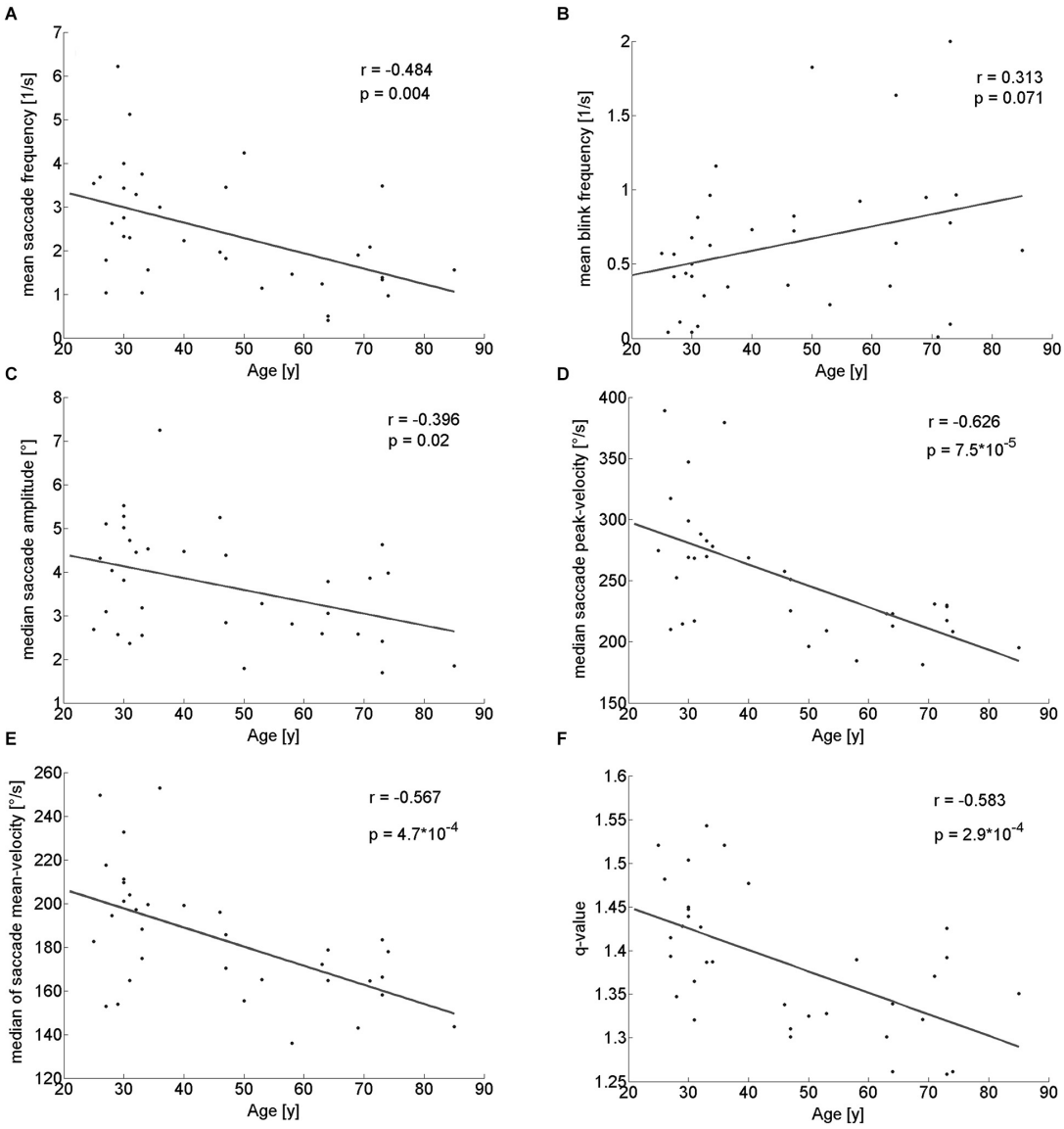

FIGURE 2 | Pearson correlation of different eye movement parameters with participants' age. There is a significant decline with age of saccade frequency **(A)**, saccade amplitude **(C)**, saccade peak- **(D)** and mean-velocity **(E)** and q-value **(F)**. Blink rate trended to be larger in older participants **(B)**.

was a negative correlation for all of them (saccade amplitude: $r_{(32)} = -0.759$; $p < 0.001$; **Figure 3A**; saccade peak-velocity: $r_{(32)} = -0.752$; $p < 0.001$; **Figure 3C**; saccade mean-velocity: $r_{(32)} = -0.731$; $p < 0.001$; **Figure 3E**; q-value: $r_{(32)} = -0.477$; $p = 0.004$; **Figure 3G**). This negative correlation remained for the variation coefficient, which serves as the standardized (normalized) measure of dispersion of the four saccadic performance parameters (saccade amplitude: $r_{(32)} = -0.349$; $p = 0.043$; **Figure 3B**; saccade peak-velocity: $r_{(32)} = -0.460$; $p = 0.006$; **Figure 3D**; saccade mean-velocity: $r_{(32)} = -0.556$; $p < 0.001$; **Figure 3F**; q-value: $r_{(32)} = -0.404$; $p = 0.018$; **Figure 3H**).

Main sequence

Saccade amplitudes and velocities are not independent from each other, but coupled through the so-called main sequence.

Table 2 | Saccade mean- and peak-velocity examined separately for each direction.

	Eye-movement parameter		U-test				
			Young (SD)	Old (SD)	*Z*-value	*p*-value	Effect-size (AUC)
Task I (walking with free gaze)	median saccade amplitude [°] for saccades	Upward	4.39 (1.46)	3.50 (1.69)	1.274	0.203	0.669 [0.484 0.855]
		Downward	5.46 (2.68)	3.79 (1.65)	1.963	0.050	0.699 [0.521 0.877]
		Leftward	7.26 (2.57)	4.40 (1.32)	3.582	0.0003	0.862 [0.734 0.990]
		Rightward	6.66 (2.29)	5.14 (2.21)	1.688	0.092	0.671 [0.489 0.854]
	median saccade mean-velocity [°/s] for saccades	Upward	147.9 (15.4)	139.8 (14.8)	0.861	0.389	0.625 [0.433 0.817]
		Downward	164.9 (32.1)	138.0 (16.9)	2.721	0.007	0.775 [0.616 0.934]
		Leftward	207.9 (41.9)	161.5 (27.2)	3.237	0.001	0.827 [0.685 0.969]
		Rightward	197.1 (32.0)	169.1 (29.3)	2.342	0.019	0.737 [0.568 0.906]
	median saccade peak-velocity [°/s] for saccades	Upward	174.8 (24.5)	170.3 (25.4)	0.241	0.810	0.559 [0.361 0.757]
		Downward	189.1 (45.5)	166.1 (21.2)	1.722	0.090	0.675 [0.493 0.857]
		Leftward	267.0 (55.4)	201.2 (38.2)	3.444	0.0006	0.848 [0.714 0.982]
		Rightward	253.4 (47.2)	199.7 (38.9)	2.961	0.003	0.799 [0.648 0.951]

Median saccade amplitude, mean- and peak-velocity of the two groups during the first task and their corresponding statistics. Only saccades with a higher mean-velocity of more than 100° for each direction were analyzed to exclude saccades with an unspecific direction.

When fitting a power function to the main sequence (see Section Methods) for these amplitude ranges (**Figure 4A**), the exponent of the power function ("L"), showed no significant difference between the groups (**Table 1**) or correlation with age ($r_{(32)} = 0.04$; $p = 0.821$; **Figure 4C**). On the other hand, the fit parameter K, which corresponds to the rise of the power function when the exponent is set, was significantly smaller in older participants (**Table 1**) and negatively correlated with age ($r_{(32)} = -0.476$; $p = 0.004$; **Figure 4B**). In addition, the analysis of saccade peak-velocity within certain amplitude ranges showed a significantly smaller peak-velocity in the older participant group for all analyzed amplitudes (**Table 1**). In sum, nearly all saccade parameters are affected by age, but the general shape of the main sequence remains remarkably unaffected.

Spontaneous tracking

During the free exploration of task I, there were periods in which participants spontaneously tracked a target during walking. These tracking movements had an average gain above 1 for nearly all (29/30) participants. There was no significant linear dependency on age ($r_{(28)} = 0.282$; $p = 0.13$; **Figure 5A**), even though a median-split analysis indicated a somewhat higher gain for the older half of observers (**Table 1**).

TASK II—TRACKING OF A STATIONARY TARGET WHILE WALKING

For the tracking task, we did not find any significant dependence of tracking gain or tracking performance (as quantified by the RMSE) on age (gain: $r_{(26)} = -0.169$; $p = 0.39$; **Figure 5B**; RMSE: $r_{(23)} = 0.064$; $p = 0.76$; **Figure 5C**). Similarly, a median-split analysis did not show any significant difference between the older and the younger half of the participants (**Table 1**).

DISCUSSION

In this study we analyzed the age-dependent changes of basic eye-movement parameters in a real-world setting during everyday tasks. Participants were free to move their eyes and head during self-motion. Some of the oculomotor parameters, such as saccade frequency and velocities, showed a significant decline with healthy aging. Others, i.e., tracking performance of an earth-fixed target during self-motion, did not appear to be influenced by age. Accordingly, our saccade data resemble most findings obtained under laboratory conditions. Our findings concerning smooth eye movements, however, challenge the transferability of eye-movement data from the laboratory to the real world.

WALKING WITH FREE GAZE

In task I, participants had to walk along a hallway with free gaze. A key result was the significant decrease of saccade peak- and mean velocity with age. Previous results on saccade peak velocity in the laboratory were inconsistent and reported both a decline with aging (Warabi et al., 1984; Sharpe and Zackon, 1987; Irving et al., 2006) as well as no significant age dependency (Henriksson et al., 1980; Munoz et al., 1998). While the data of Henriksson et al. (1980) might have had too little statistical power to show a significant effect (6–7 participants per age-group and about 10 saccades for each amplitude investigated), Munoz et al. (1998) investigated only saccades of an amplitude of 20°. Others have hypothesized that an amplitude-dependent saturation in saccade peak velocity in older participants might only affect saccade velocities for amplitudes exceeding 20° (Moschner and Baloh, 1994), which could not be confirmed by our results. The main sequence fit-parameters showed a continuous, significant decline of saccade peak velocity with age.

Mean saccade velocity in the elderly has rarely been examined before. Our results showed a clear reduction for older participants, which was also reported by Spooner et al. (1980) and, with a less strong reduction, by Abel et al. (1983). This effect might be explained by the results of Munoz et al. (1998), who found a significant increase in saccade duration in the oldest participants without an accompanying drop in peak velocity for saccades of the same amplitude. This finding implies a change of saccade skewness with age. Indeed,

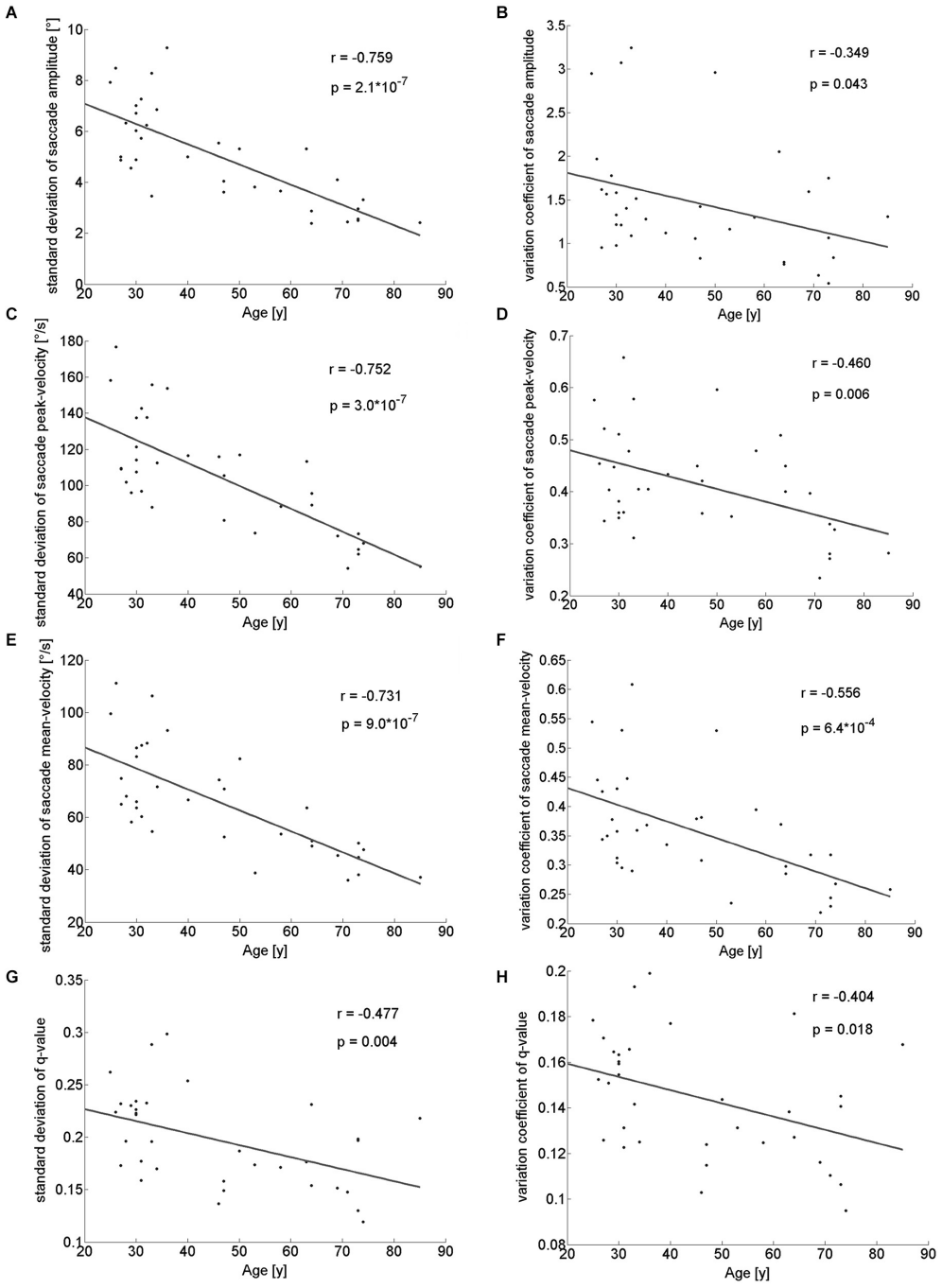

FIGURE 3 | Standard deviations and variation coefficient of saccade amplitude (A,B), peak- (C,D) and mean-velocities (E,F) and q-value (G,H) in relation to the age of the participants. There is a clearly negative correlation for these parameters indicating a decreased variability of saccadic performance in older participants.

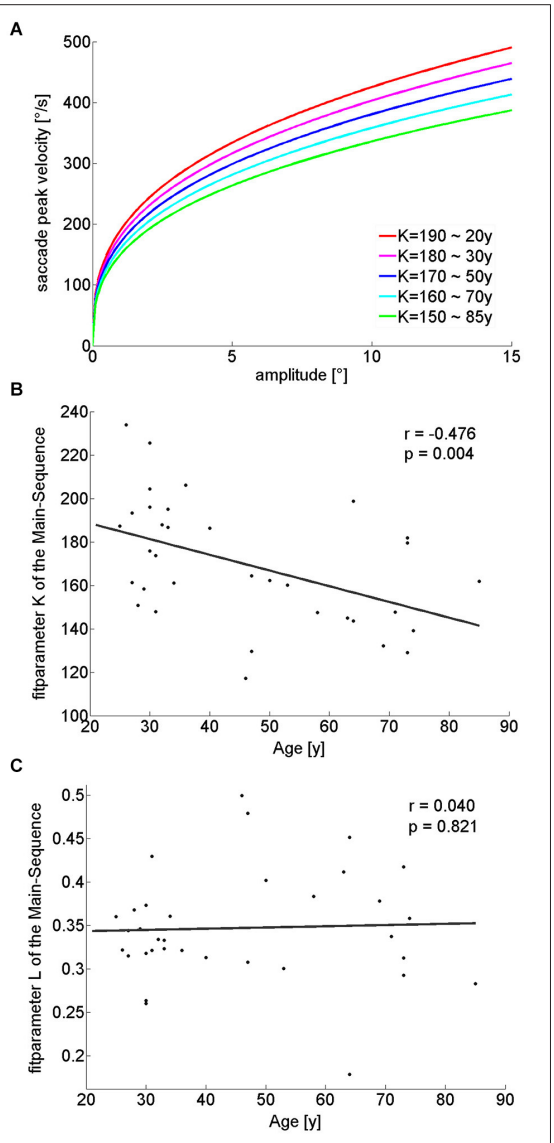

FIGURE 4 | Illustration of Main-sequence power-function fits for different K-values, while the exponent ("L") remains unchanged (A), each K-value relates to a different age cohort, as extracted from the linear regression of K-value over age (B). L-values were constant at 0.35 for all fits **(C)**. With increasing age, smaller K-values **(B)** lead to decreasing slopes of the power functions and therefore smaller saccade peak-velocities for the same amplitudes when L-values are constant **(C)**.

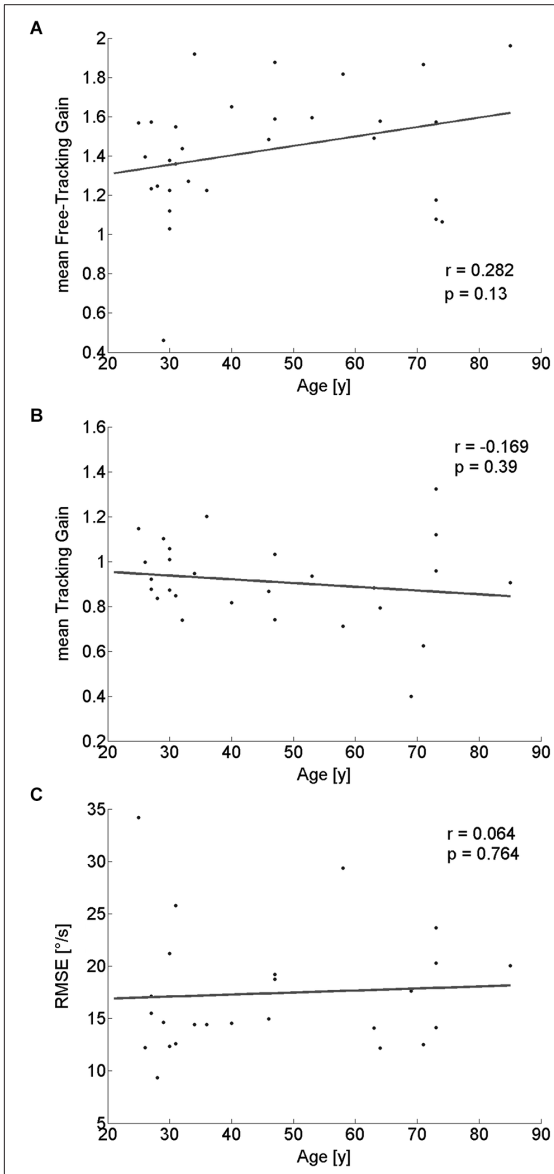

FIGURE 5 | Pearson correlation of tracking-performance parameters with participants' age. Neither tracking gain during free exploration **(A)**, nor gain **(B)** or RMSE **(C)** of the tracking eye-movement of a specific target showed a significant correlation with age.

a generally altered saccade velocity profile of the elderly is also suggested by the q-values as observed in our study. Older participants showed significantly smaller q-values, indicating a less curved saccade velocity profile. This is reflected by a less increased saccade peak-velocity as compared to the saccade mean-velocity.

Another interesting aspect is the separately analyzed age-related difference between horizontal and vertical saccades. For vertical saccades, Huaman and Sharpe (1993) previously reported a decrease in the maximal voluntary excursion in the elderly but no significantly different saccade peak velocity. Our results show a different picture, since saccade amplitude and

mean- and peak-velocities were significantly smaller in the older group especially in the horizontal plane. In contrast, in the vertical domain, only the downward saccades showed a trend for being smaller in older participants. This might be mainly due to the nature of the hallway, in which the ceiling is comparably uninteresting/uninformative, as objects and other potential targets of exploration were mostly present in the horizontal periphery.

The general decrease of saccade speed could be due to a loss of contractibility (McKelvie et al., 1999) or mechanical efficiency (Clark and Demer, 2002) of the eye muscles while aging. Concerning their neural basis, saccade properties are defined by burst neurons in the paramedian pontine reticular formation and are not under voluntary control (Sparks, 2002). The significantly decreased saccade velocities and the smaller q-values for older participants suggest a generally lower activity and a less marked burstiness of these neurons. While the brainstem itself seems to be unaffected by healthy aging (Brody and Vijayashankar, 1977), these neurons receive input from the frontal eye fields, superior colliculus, parietal cortex, and basal ganglia (Wurtz and Goldberg, 1989; Sparks, 2002; Leigh and Zee, 2006). Reduced function in one of these areas or brain regions in older participants could be responsible for a decreased firing frequency of premotor or motor neurons and therefore the reduced saccade velocities. For example, it has been shown that frontal lobe lesions can lead to a slowing of saccades (Sharpe, 1986). Indeed, there are studies suggesting a prefrontal functional (Fabiani and Friedman, 1997) and structural (West, 1996) decline with aging. Additionally some studies showed a decreased neuronal density (Huttenlocher, 1979) and a loss of cortical gray matter in the elderly (Pfefferbaum et al., 1994), which might also contribute to a decrease in saccade velocity.

The reduction of saccade frequency, amplitude and velocities could be attributed to a more narrow viewing area of elderly people. This is in line with the decreased variability of saccadic performance in older participants, as shown by the negative correlation of standard deviation of saccadic parameters with age in our study. Yet, the negative correlation of the variation coefficient with age shows, that saccade performance is in general less variable in the elderly. A more narrow viewing area in the elderly is also supported by our results on the separately analyzed horizontal and vertical saccades, which show a consistent decrease in saccade amplitude and velocities in the older group, especially for saccades in the horizontal plane. This might be due to higher effort while walking, e.g., more looking on the pathway to avoid obstacles and plan appropriate motor responses (Di Fabio et al., 2003; 't Hart and Einhäuser, 2012), or less confidence in exploring while walking due to a higher likelihood and cost of a potential fall (Hadley et al., 1985). This is in line with the structure of the hallway, in which exploration targets were mostly present in the horizontal periphery. Since objects are the dominant driver of fixations (Stoll et al., 2015), less exploration might be the main reason for the significantly decreased horizontal saccade amplitudes and velocities in the older participants. Chapman and Hollands (2006) have shown that older adults looked significantly earlier to targets, and fixated the targets for longer periods than younger adults while walking along a pathway. The authors explained their result as a consequence of age-related decline in general visual function (Morgan, 1993), slowed cognitive processing (Salthouse, 1996) and decline in visuomotor processing (Moschner and Baloh, 1994).

TRACKING EYE-MOVEMENTS DURING SELF-MOTION

In our study, the performance of tracking eye-movements while walking showed different results for spontaneous tracking movements (task I) as compared to instructed tracking movements (task II). Active tracking of optic flow elements in the laboratory has been reported to have a gain close to perfect, i.e., 1.0 (Niemann et al., 1999), which was also the case in our study during tracking of a given stationary target. On the other hand, the gain of most participants during spontaneous tracking was greater than 1.0. This result could be due to the fact that objects in the real world, unlike most artificial stimuli and the target used in our study, have a considerable extent. This leads to more eye movements across the object during the tracking and eventually to a higher speed of the eye during tracking. On the other hand, the computation of optic flow fields of real world scenes can be imprecise because of light reflections or plain surfaces (Gautama and Van Hulle, 2002). This might have led to an underestimation of target velocity in this task. Nevertheless, a median-split analysis showed that the gain of the group of older participants was significantly higher as compared to the younger participants. This suggests a generally more imprecise tracking eye-movement of freely-chosen targets during walking in the elderly. One possible explanation could be an age-related, gradually functional decline of the visual-vestibular system. It has been shown in the laboratory that visual influences on the vestibulo-ocular reflex decline in the elderly together with a deterioration of visual following (Paige, 1994). Accordingly, tracking eye-movements in the elderly might get affected due to differing available input signals.

Unlike smooth pursuit in the laboratory, which has been shown to get worse in the elderly as compared to younger adults (Moschner and Baloh, 1994; Ross et al., 1999), visual tracking performance of a given target during self-motion appeared to be unaffected by age in our study. This finding might suggest compensatory mechanisms, e.g., head movements, or additional sensory cues like optic flow or vestibular signals, which help to maintain normal performance. Paige (1994) found an increased likelihood and intensity of circular vection, a psychophysical measure of visual-vestibular interactions, and proposed an enhanced perception of self-motion in the elderly, which might serve as a visual compensation for age-dependent loss of vestibular cues. Accordingly, optic flow information could neuronally be given a stronger weight in the process of eye-movement control during self-motion. Such an enhanced weight of sensory self-motion information could explain the decreased smooth-pursuit gain in the laboratory due to its absence when measuring with a restrained head. On the other hand motion perception and detection of random dot patterns in the laboratory have been shown to deteriorate in the elderly (Tran et al., 1998). Along similar lines, Billino et al. (2008)

showed a gradual decrease in the perception of two-dimensional translational motion and biological motion in the elderly. In contrast, heading detection via expanding radial flow fields was stable across the lifespan in this study (Billino et al., 2008). Nevertheless, a recent study of Lich and Bremmer (2014) showed a decreased absolute heading performance in the elderly in a virtual-reality setting. In this study, the authors were able to model their results in a neural network of visual self-motion processing by an age related neuronal cell loss in area MST. Taken together, this suggests an impairment of motion-selective areas in the brain, such as the middle temporal (MT) area (Newsome and Paré, 1988), the MST area (Duffy and Wurtz, 1991; Bremmer et al., 1999) and the VIP area, which is particularly important in decoding global motion and heading information (Bremmer et al., 2002a; Bremmer, 2005; Chen et al., 2011).

The suggested relevance of self-motion processing for oculomotor performance in the real world as compared to laboratory settings is supported by studies in schizophrenia patients. These patients show an impaired smooth pursuit in the laboratory (Holzman et al., 1974; O'Driscoll and Callahan, 2008) but only a subtle change of tracking eye-movement performance in the real world (Dowiasch et al., 2014). The importance of additional sensory signals to the visual system of schizophrenia patients has been shown by a study of Holzman (2000), in which patients performed worse in a velocity discrimination task when additional non-velocity stimulus cues were eliminated. In natural behavior, these additional cues are almost always present and might serve as a support or even substitute to compensate impairments of specific visual functions. The tracking of a moving object in a real-world situation, while participants are not moving but able to move their head, is an interesting issue for future research. Such a paradigm might be more closely linked to smooth pursuit in the laboratory and therefore might bridge the gap between the results in the literature and the real-world data in our study.

Being aware of the numerous differences between our real-world tasks and typical laboratory measurements, our results provide a first step towards analyzing real-world oculomotor behavior in healthy aging. Furthermore, our correlation analysis together with the effect sizes allowed us to examine to what extent age influences different eye-movement parameters in real-world situations. Together with two recently published studies (Marx et al., 2012; Dowiasch et al., 2014) our results highlight the possible advantages of mobile eye tracking as a fast, reliable, objective and easy-to-use tool, especially when investigating clinical populations or the elderly. Identifying the sources of differences and commonalities between laboratory results and real-world data will be an important issue for future research. Especially the multimodal area VIP (Bremmer et al., 2002b; Schlack et al., 2005), for which a functional equivalent has been identified in humans (Bremmer et al., 2001), might play a crucial role in natural contexts by providing and combining additional sensory information. How such key areas in motion processing are related to the changing oculomotor behavior during aging, and how they integrate the rich information available in the real world for gaze control remains an exciting topic for further research. In any case, our present study underlines the need for addressing real-world situations to fully understand the impact of neuronal changes on oculomotor function and motor behavior in general during healthy aging.

ACKNOWLEDGMENTS

This work was supported by the "Deutsche Forschungsgemeinschaft" (DFG) through the International Research Training Group IRTG 1901 "The Brain in Action—BrainAct", the Collaborative Research Center SFB/TRR 135 "Cardinal Mechanisms of Perception" and grant EI 852/3.

REFERENCES

Abel, L. A., Troost, B. T., and Dell'Osso, L. F. (1983). The effects of age on normal saccadic characteristics and their variability. *Vision Res.* 23, 33–37. doi: 10.1016/0042-6989(83)90038-x

Bahill, A. T., Clark, M. R., and Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Math. Biosci.* 24, 191–204. doi: 10.1016/0025-5564(75)90075-9

Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *J. Math. Psychol.* 12, 387–415. doi: 10.1016/0022-2496(75)90001-2

Billino, J., Bremmer, F., and Gegenfurtner, K. R. (2008). Differential aging of motion processing mechanisms: evidence against general perceptual decline. *Vision Res.* 48, 1254–1261. doi: 10.1016/j.visres.2008.02.014

Bremmer, F. (2005). Navigation in space—the role of the macaque ventral intraparietal area. *J. Physiol.* 566, 29–35. doi: 10.1113/jphysiol.2005.082552

Bremmer, F., Duhamel, J.-R., Ben Hamed, S., and Graf, W. (2002a). Heading encoding in the macaque ventral intraparietal area (VIP). *Eur. J. Neurosci.* 16, 1554–1568. doi: 10.1046/j.1460-9568.2002.02207.x

Bremmer, F., Duhamel, J.-R., Hamed, S. B., and Graf, W. (2000). Stages of self-motion processing in primate posterior parietal cortex. *Int. Rev. Neurobiol.* 44, 173–198. doi: 10.1016/S0074-7742(08)60742-4

Bremmer, F., Klam, F., Duhamel, J.-R., Ben Hamed, S., and Graf, W. (2002b). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Eur. J. Neurosci.* 16, 1569–1586. doi: 10.1046/j.1460-9568.2002.02206.x

Bremmer, F., Kubischik, M., Pekel, M., Lappe, M., and Hoffmann, K.-P. (1999). Linear vestibular self-motion signals in monkey medial superior temporal area. *Ann. N Y Acad. Sci.* 871, 272–281. doi: 10.1111/j.1749-6632.1999.tb09191.x

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296. doi: 10.1016/s0896-6273(01)00198-2

Britten, K. H. (2008). Mechanisms of self-motion perception. *Annu. Rev. Neurosci.* 31, 389–410. doi: 10.1146/annurev.neuro.29.051605.112953

Brody, H., and Vijayashankar, N. (1977). "Anatomical changes in the nervous system," in *The Biology of Aging*, eds C. E. Finch and L. Hayflick (New York: van Nostrand Reinhold), 241–261.

Chapman, G. J., and Hollands, M. A. (2006). Evidence for a link between changes to gaze behaviour and risk of falling in older adults during adaptive locomotion. *Gait Posture* 24, 288–294. doi: 10.1016/j.gaitpost.2005.10.002

Chen, A., DeAngelis, G. C., and Angelaki, D. E. (2011). Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *J. Neurosci.* 31, 12036–12052. doi: 10.1523/JNEUROSCI.0395-11.2011

Clark, R. A., and Demer, J. L. (2002). Effect of aging on human rectus extraocular muscle paths demonstrated by magnetic resonance imaging. *Am. J. Ophthalmol.* 134, 872–878. doi: 10.1016/s0002-9394(02)01695-1

Di Fabio, R. P., Greany, J. F., and Zampieri, C. (2003). Saccade-stepping interactions revise the motor plan for obstacle avoidance. *J. Mot. Behav.* 35, 383–397. doi: 10.1080/00222890309603158

Dowiasch, S., Backasch, B., Einhäuser, W., Leube, D., Kircher, T., and Bremmer, F. (2014). Eye movements of patients with schizophrenia in a natural environment. *Eur. Arch. Psychiatry Clin. Neurosci.* doi: 10.1007/s00406-014-0567-8. [Epub ahead of print].

Duffy, C. J., and Wurtz, R. H. (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J. Neurophysiol.* 65, 1329–1345.

Fabiani, M., and Friedman, D. (1997). Dissociations between memory for temporal order and recognition memory in aging. *Neuropsychologia* 35, 129–141. doi: 10.1016/s0028-3932(96)00073-5

Foulsham, T., Walker, E., and Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Res.* 51, 1920–1931. doi: 10.1016/j.visres.2011.07.002

Gautama, T., and Van Hulle, M. M. (2002). A phase-based approach to the estimation of the optical flow field using spatial filtering. *IEEE Trans. Neural Netw.* 13, 1127–1136. doi: 10.1109/TNN.2002.1031944

Gooding, D. C., and Basso, M. A. (2008). The tell-tale tasks: a review of saccadic research in psychiatric patient populations. *Brain Cogn.* 68, 371–390. doi: 10.1016/j.bandc.2008.08.024

Gooding, D. C., Iacono, W. G., and Beiser, M. (1994). Temporal stability of smooth-pursuit eye tracking in first-episode psychosis. *Psychophysiology* 31, 62–67. doi: 10.1111/j.1469-8986.1994.tb01025.x

Gu, Y., Angelaki, D. E., and Deangelis, G. C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* 11, 1201–1210. doi: 10.1038/nn.2191

Hadley, E., Radebaugh, T. S., and Suzman, R. (1985). Falls and gait disorders among the elderly. A challenge for research. *Clin. Geriatr. Med.* 1, 497–500.

Hanley, J. A., and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36. doi: 10.1148/radiology.143.1.7063747

Hayhoe, M., and Ballard, D. (2005). Eye movements in natural behavior. *Trends Cogn. Sci.* 9, 188–194. doi: 10.1016/j.tics.2005.02.009

Henriksson, N. G., Pyykkö, I., Schalen, L., and Wennmo, C. (1980). Velocity patterns of rapid eye movements. *Acta Otolaryngol.* 89, 504–512. doi: 10.3109/00016488009127168

Hentschke, H., and Stüttgen, M. C. (2011). Computation of measures of effect size for neuroscience data sets. *Eur. J. Neurosci.* 34, 1887–1894. doi: 10.1111/j.1460-9568.2011.07902.x

Holzman, P. S. (2000). Eye movements and the search for the essence of schizophrenia. *Brain Res. Brain Res. Rev.* 31, 350–356. doi: 10.1016/s0165-0173(99)00051-x

Holzman, P. S., Proctor, L. R., Levy, D. L., Yasillo, N. J., Meltzer, H. Y., and Hurt, S. W. (1974). Eye-tracking dysfunctions in schizophrenic patients and their relatives. *Arch. Gen. Psychiatry* 31, 143–151. doi: 10.1001/archpsyc.1974.01760140005001

Huaman, A. G., and Sharpe, J. A. (1993). Vertical saccades in senescence. *Invest. Ophthalmol. Vis. Sci.* 34, 2588–2595.

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex — developmental changes and effects of aging. *Brain Res.* 163, 195–205. doi: 10.1016/0006-8993(79)90349-4

Inchingolo, P., Spanio, M., and Bianchi, M. (1987). "The characteristic peak velocity-mean velocity of saccadic eye movements in man," in *Eye Movements: From Physiology to Cognition*, eds J. K. O'Regan and A. Lévy-Schoen (Amsterdam: Elsevier Science Publishers), 17–26.

Irving, E. L., Steinbach, M. J., Lillakas, L., Babu, R. J., and Hutchings, N. (2006). Horizontal saccade dynamics across the human life span. *Invest. Ophthalmol. Vis. Sci.* 47, 2478–2484. doi: 10.1167/iovs.05-1311

Klein, C., Fischer, B., Hartnegg, K., Heiss, W. H., and Roth, M. (2000). Optomotor and neuropsychological performance in old age. *Exp. Brain Res.* 135, 141–154. doi: 10.1007/s002210000506

Land, M., Mennie, N., and Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception* 28, 1311–1328. doi: 10.1068/p2935

Leigh, R. J., and Zee, D. S. (2006). *The Neurology of Eye Movements.* (Bd. 90). New York: Oxford University Press.

Lich, M., and Bremmer, F. (2014). Self-Motion perception in the elderly. *Front. Hum. Neurosci.* 8:681. doi: 10.3389/fnhum.2014.00681

Mann, H. B., and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18, 50–60. doi: 10.1214/aoms/1177730491

Marx, S., Respondek, G., Stamelou, M., Dowiasch, S., Stoll, J., Bremmer, F., et al. (2012). Validation of mobile eye-tracking as novel and efficient means for differentiating progressive supranuclear palsy from Parkinson's disease. *Front. Behav. Neurosci.* 6:88. doi: 10.3389/fnbeh.2012.00088

McKelvie, P., Friling, R., Davey, K., and Kowal, L. (1999). Changes as the result of ageing in extraocular muscles: a post-mortem study. *Aust. N. Z. J. Ophthalmol.* 27, 420–425. doi: 10.1046/j.1440-1606.1999.00244.x

Morgan, M. W. (1993). "Normal age related vision changes," in *Vision and Aging*, eds A. A. Rosenbloom and M. Morgan 2nd Edn. (Boston: Butterworth-Heinemann), 178–199.

Moschner, C., and Baloh, R. W. (1994). Age-Related changes in visual tracking. *J. Gerontol.* 49, M235–M238. doi: 10.1093/geronj/49.5.m235

Munoz, D. P., Broughton, J. R., Goldring, J. E., and Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Exp. Brain Res.* 121, 391–400. doi: 10.1007/s002210050473

Newsome, W. T., and Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* 8, 2201–2211.

Niemann, T., Lappe, M., Büscher, A., and Hoffmann, K.-P. (1999). Ocular responses to radial optic flow and single accelerated targets in humans. *Vision Res.* 39, 1359–1371. doi: 10.1016/s0042-6989(98)00236-3

O'Driscoll, G. A., and Callahan, B. L. (2008). Smooth pursuit in schizophrenia: a meta-analytic review of research since 1993. *Brain Cogn.* 68, 359–370. doi: 10.1016/j.bandc.2008.08.023

Paige, G. D. (1994). Senescence of human visual-vestibular interactions: smooth pursuit, optokinetic and vestibular control of eye movements with aging. *Exp. Brain Res.* 98, 355–372. doi: 10.1007/bf00228423

Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V., Rawles, J. M., Zipursky, R. B., and Lim, K. O. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch. Neurol.* 51, 874–887. doi: 10.1001/archneur.1994.00540210046012

Pinkhardt, E. H., Jürgens, R., Becker, W., Valdarno, F., Ludolph, P. D. A. C., and Kassubek, P. D. J. (2008). Differential diagnostic value of eye movement recording in PSP-parkinsonism, Richardson's syndrome and idiopathic Parkinson's disease. *J. Neurol.* 255, 1916–1925. doi: 10.1007/s00415-009-0027-y

Pitzalis, S., Sdoia, S., Bultrini, A., Committeri, G., Di Russo, F., Fattori, P., et al. (2013). Selectivity to translational egomotion in human brain motion areas. *PLoS One* 8:e60241. doi: 10.1371/journal.pone.0060241

Ross, R. G., Olincy, A., Harris, J. G., Radant, A., Adler, L. E., Compagnon, N., et al. (1999). The effects of age on a smooth pursuit tracking task in adults with schizophrenia and normal subjects. *Biol. Psychiatry* 46, 383–391. doi: 10.1016/s0006-3223(98)00369-2

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037//0033-295x.103.3.403

Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K. P., and Bremmer, F. (2005). Multisensory space representations in the macaque ventral intraparietal area. *J. Neurosci.* 25, 4616–4625. doi: 10.1523/jneurosci.0455-05.2005

Schneider, E., Villgrattner, T., Vockeroth, J., Bartl, K., Kohlbecher, S., Bardins, S., et al. (2009). EyeSeeCam: an eye movement-driven head camera for the examination of natural visual exploration. *Ann. N Y Acad. Sci.* 1164, 461–467. doi: 10.1111/j.1749-6632.2009.03858.x

Sharpe, J. A. (1986). "Adaptation to frontal lobe lesions," in *Adaptive Processes in Visual and Oculomotor Systems*, eds E. L. Keller and D. S. Zee (Oxford: Pergamon Press), 239–246.

Sharpe, J. A., and Zackon, D. H. (1987). Senescent saccades: effects of aging on their accuracy, latency and velocity. *Acta Otolaryngol.* 104, 422–428. doi: 10.3109/00016488709128270

Smyrnis, N. (2008). Metric issues in the study of eye movements in psychiatry. *Brain Cogn.* 68, 341–358. doi: 10.1016/j.bandc.2008.08.022

Sparks, D. L. (2002). The brainstem control of saccadic eye movements. *Nat. Rev. Neurosci.* 3, 952–964. doi: 10.1038/nrn986

Spooner, J. W., Sakala, S. M., and Baloh, R. W. (1980). EFfect of aging on eye tracking. *Arch. Neurol.* 37, 575–576. doi: 10.1001/archneur.1980.00500580071012

Stoll, J., Thrun, M., Nuthmann, A., and Einhäuser, W. (2015). Overt attention in natural scenes: objects dominate features. *Vision Res.* 107, 36–48. doi: 10.1016/j.visres.2014.11.006.

Tatler, B. W., Hayhoe, M. M., Land, M. F., and Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *J. Vis.* 11:5. doi: 10.1167/11.5.5

't Hart, B. M., and Einhäuser, W. (2012). Mind the step: complementary effects of an implicit task on eye and head movements in real-life gaze allocation. *Exp. Brain Res.* 223, 233–249. doi: 10.1007/s00221-012-3254-x

't Hart, B. M., Vockeroth, J., Schumann, F., Bartl, K., Schneider, E., König, P., et al. (2009). Gaze allocation in natural stimuli: comparing free exploration to head-fixed viewing conditions. *Visual Cogn.* 17, 1132–1158. doi: 10.1080/13506280902812304

Tran, D. B., Silverman, S. E., Zimmerman, K., and Feldon, S. E. (1998). Age-related deterioration of motion perception and detection. *Graefes Arch. Clin. Exp. Ophthalmol.* 236, 269–273. doi: 10.1007/s0041700 50076

Wall, M. B., and Smith, A. T. (2008). The representation of egomotion in the human brain. *Curr. Biol.* 18, 191–194. doi: 10.1016/j.cub.2007.12.053

Warabi, T., Kase, M., and Kato, T. (1984). Effect of aging on the accuracy of visually guided saccadic eye movement. *Ann. Neurol.* 16, 449–454. doi: 10.1002/ana.410160405

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychol. Bull.* 120, 272–292. doi: 10.1037//0033-2909.120.2.272

Wurtz, R. H., and Goldberg, M. E. (1989). *The Neurobiology of Saccadic Eye Movements.* New York: Elsevier Publishing Company.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 October 2014; accepted: 17 January 2015; published online: 10 February 2015.

Citation: Dowiasch S, Marx S, Einhäuser W and Bremmer F (2015) Effects of aging on eye movements in the real world. Front. Hum. Neurosci. ***9****:46. doi: 10.3389/fnhum.2015.00046*

This article was submitted to the journal Frontiers in Human Neuroscience.

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 13 December 2012
doi: 10.3389/fnbeh.2012.00088

Validation of mobile eye-tracking as novel and efficient means for differentiating progressive supranuclear palsy from Parkinson's disease

Svenja Marx[1]*†, Gesine Respondek[2,3]†, Maria Stamelou[2,4], Stefan Dowiasch[1], Josef Stoll[1], Frank Bremmer[1], Wolfgang H. Oertel[2], Günter U. Höglinger[2,3,5]‡ and Wolfgang Einhäuser[1,6]‡

[1] Department of Neurophysics, Philipps-University, Marburg, Germany
[2] Department of Neurology, Philipps-University, Marburg, Germany
[3] Department of Neurology, Technische Universität München, Munich, Germany
[4] Sobell Department for Motor Neurosciences and Movement Disorders, Institute of Neurology, University College London, London, UK
[5] German Center for Neurodegenerative Diseases (DZNE), Munich, Germany
[6] Center for Interdisciplinary Research, Bielefeld University, Bielefeld, Germany

Edited by:
Markus Lappe, Universität Münster, Germany

Reviewed by:
Florian Ostendorf, Charité - Universitätsmedizin Berlin, Germany
Uwe Ilg, Hertie-Institute for Clinical Brain Research, Germany

***Correspondence:**
Svenja Marx, Department of Neurophysics, Philipps-University, AG Neurophysik, Karl-von-Frisch Straße 8a, 35032 Marburg, Germany.
e-mail: svenja.marx@physik.uni-marburg.de
†,‡These authors equally contributed to this work.

Background: The decreased ability to carry out vertical saccades is a key symptom of Progressive Supranuclear Palsy (PSP). Objective measurement devices can help to reliably detect subtle eye movement disturbances to improve sensitivity and specificity of the clinical diagnosis. The present study aims at transferring findings from restricted stationary video-oculography (VOG) to a wearable head-mounted device, which can be readily applied in clinical practice. **Methods:** We investigated the eye movements in 10 possible or probable PSP patients, 11 Parkinson's disease (PD) patients, and 10 age-matched healthy controls (HCs) using a mobile, gaze-driven video camera setup (EyeSeeCam). Ocular movements were analyzed during a standardized fixation protocol and in an unrestricted real-life scenario while walking along a corridor. **Results:** The EyeSeeCam detected prominent impairment of both saccade velocity and amplitude in PSP patients, differentiating them from PD and HCs. Differences were particularly evident for saccades in the vertical plane, and stronger for saccades than for other eye movements. Differences were more pronounced during the standardized protocol than in the real-life scenario. **Conclusions:** Combined analysis of saccade velocity and saccade amplitude during the fixation protocol with the EyeSeeCam provides a simple, rapid (<20 s), and reliable tool to differentiate clinically established PSP patients from PD and HCs. As such, our findings prepare the ground for using wearable eye-tracking in patients with uncertain diagnoses.

Keywords: progressive supranuclear palsy, mobile eye-tracking, eye movements, Parkinson's disease, video-oculography

INTRODUCTION

Eye movement abnormalities are an essential clinical feature of Progressive Supranuclear Palsy (PSP). Vertical supranuclear gaze palsy or decreased velocities of vertical saccades are a key to the clinical diagnosis of PSP (Litvan et al., 1996). Besides their role as diagnostic signs, eye movement abnormalities disable PSP patients in their daily routine.

Stationary video-oculography (VOG) during head-fixed viewing shows that virtually all forms of eye movements are affected in PSP, with saccadic eye movements being most prominently impaired. Particularly vertical saccades show reduced amplitude and peak velocity when compared to Parkinson's disease (PD) patients and healthy controls (HCs) (Pinkhardt et al., 2008; Chen et al., 2010; Pinkhardt and Kassubek, 2011). Vergence movements and the associated modulation of the linear vestibuloocular reflex are also considerably affected (Chen et al., 2010). The presence of horizontal square wave jerks during attempted fixation of stationary targets is characteristic of PSP (Chen et al., 2010; Otero-Millan et al., 2011). Among these deficits, saccadic peak velocity in the vertical plane shows the sharpest contrast between PSP and PD (Pinkhardt and Kassubek, 2011).

These PSP-specific eye movement abnormalities make clinical investigation of eye movements in patients with Parkinsonian syndromes of great value for differential diagnosis. Correct diagnosis of PSP remains challenging, especially in its early stages (Burn and Lees, 2002). Eye movement abnormalities are not always easy to detect clinically. Particularly, slowing of saccades is a characteristic symptom that can be missed by less experienced neurologists.

Objective measurement devices aid detection of subtle eye movement disturbances. Stationary VOG setups typically require careful calibration, need patient collaboration, and are thus largely impractical for clinical routine. Head-fixed viewing lacks vestibular and other cross-modal information, leaving the relevance of observed eye movement impairment for real-life behavior open. As a first step toward the development of an

objective, easy-to-use method for eye movement-based diagnosis, we here tested if recording eye movements with the versatile, head-mounted EyeSeeCam (Brandt et al., 2006; Schneider et al., 2006, 2009) in a brief and simple fixation protocol can differentiate between patients with clinically established PSP as compared to established PD and HCs, and measured gaze in these groups during free behavior. We aimed at establishing the EyeSeeCam's usage in PD and PSP cases and validating its discriminative power between these groups. The parameters established in the present study in clinically established patients shall pave the way for prospective studies with uncertain diagnoses.

MATERIALS AND METHODS

PARTICIPANTS

Patients examined in the Department of Neurology of the University of Marburg qualified for participation in the study, if they had clinically possible or probable PSP (Litvan et al., 1996) and were not more advanced than Hoehn and Yahr stage IV (Golbe and Ohman-Strickland, 2007). As defined by the NINDS-SPSP criteria (Litvan et al., 1996), all patients had supranuclear gaze palsy or slowing of vertical saccades at the time of examination, as evidenced by an examiner specialized in the clinical evaluation of ocular movements.

As controls, we included patients with clinically probable PD (Gibb and Lees, 1988) and HCs. HCs were free of neurologic, systemic, or psychiatric diseases, including alcohol or substance abuse, as verified by detailed evaluation of their medical histories and a comprehensive physical examination.

Further exclusion criteria were other neurological disorders, dementia (mini mental status examination <24), presently active psychiatric disorder (e.g., depression or psychosis), structural brain lesion (e.g., brain surgery, stroke with persistent neurological deficit), cataract, or other neuro-ophthalmological disorders leading to functionally relevant impairment. Since glasses cannot be worn with the EyeSeeCam, people requiring visual correction by glasses stronger than ±2 dpt were also excluded.

Before inclusion into the study, participants gave their informed written consent. All procedures conformed to the Declaration of Helsinki and were approved by the local ethics committee (Ethikkommission FB20, Philipps-Universität Marburg).

EYE AND HEAD MOVEMENT RECORDINGS

We used a mobile VOG setup (EyeSeeCam) to record the participants' eye and head movements. Participants accustomed themselves to wearing the device, while the experimental procedure was explained.

The head-mounted device consists of a head-fixed camera to record the perspective of the head, two high-speed cameras tracking eye-in-head movements, and a camera, which is automatically aligned with the observer's direction of gaze. Gaze- and head-centered videos are recorded at 25 Hz (**Figure 1A**; Movie 1 in supplementary material); eye movements at 300 Hz.

According to manufacturer's specifications, the spatial resolution of the eye-tracking device is given to 0.02° and the precision

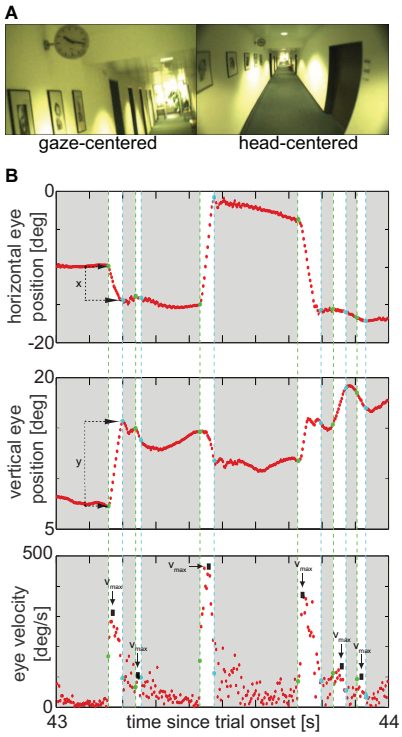

FIGURE 1 | (A) Example frame at 43.81 s in the real-life measurement, while a PD patient was looking at the clock; left: gaze camera, right: head camera. The movie of this scene including velocity histograms is shown as supplemental online Movie 1. **(B)** Eye-traces of the scene. Upper panel: Horizontal eye position, indicating the horizontal amplitude of saccades; starting and end points of saccades are marked by green and cyan dashes lines, respectively; durations of saccades are highlighted by a white background. Middle panel: Vertical eye position, indicating the vertical amplitude of the same saccades. Lower panel: Absolute eye velocity, arrows mark saccade peak velocities, used for analysis.

(relative error) on the order of 0.1° ("maximal resolution error," Schneider et al., 2009). The accuracy (absolute error) of the device under ideal conditions is about 0.5° according to specifications, and can substantially worsen if the goggles move relative to the head during prolonged measurements without recalibration. Hence, all analysis reported here only use relative measures, which are unaffected by these drifts, such as velocities and saccade amplitudes.

Being not concerned with absolute gaze orientation (i.e., with high accuracy) comes at the advantage that the device may be operated using an internal ("default") model of ocular geometry for all participants. In this mode of operation, the mapping from eye measurements on gaze direction does not require a subject-specific calibration, which is in particular beneficial in patients with limited ocular motor control or limited compliance with instructions. Although this sacrifices some precision (depending on the actual head shape compared to the default

model), no systematic effect on the measures analyzed here can be expected. For the "fixation protocol" (see below), the default model was used in all participants; for the "real-life measurements" (see below), in those participants, in whom it was possible, the subject-specific model obtained from the fixation protocol was used; for the remainder the default model was also used in real-life measurements. Since the subject-specific adaptation of the model represents a calibration procedure for absolute position, for the real-life measurement, these participants will be referred to as successfully and unsuccessfully calibrated, respectively.

When extracting head movements from the head fixed camera, for the analysis conducted here, the spatial resolution is limited by the pixel width of about 0.3°, even though sub-pixel analysis would be possible in principle. When analysis is based on subsequent frames, this limits the resolution for head movements to about 7.5°/s. While integration over multiple frames would be possible to lower this number, this would come at the cost of lower temporal resolution and thus possibly lumping distinct head movements into one.

Fixation protocol

To test the utility of the EyeSeeCam as diagnostic tool, we employed a "fixation protocol." In addition to being the first experimental part, this protocol also served to refine the EyeSeeCam's calibration for the subsequent real-life experiments by adapting the system's internal eye model to the individual. During the fixation protocol, the participants' heads were unrestrained, but they were asked to avoid head movements as far as possible. They were instructed to move their eyes to look successively at 5 laser dots projected onto a wall straight ahead, a central dot and four at 8.5° in the cardinal directions. An experimenter pointed with a finger at the dot the participant should look at. To give the participant the possibility to self-pace their fixations, presentation of the dots in time was to some degree flexible and not exactly clocked. However, the participant had to look at each dot for 2 s at least once in a time span of approximately 20 s. While this procedure is far less constrained and standardized than usual laboratory measurements, it is still more controlled than the real-life conditions of the present study. This flexible and efficient procedure makes the participation of very severely affected patients possible, presenting a clear advantage over more constrained settings.

Real-life behavior

For measuring a large range of gaze behaviors as occurring in real-life situations, we asked participants to perform a series of tasks, while spontaneous eye and head movements were recorded. First, free-exploration behavior was assessed by asking participants to walk along a 50 m corridor. Right before the participant turned around at the end of the corridor, an experimenter laid two paper spots on the floor to assess tracking behavior. Participants were asked to track the dots with their eyes, while walking back toward them. Finally, participants took the elevator and descended one-level to test a situation without active movement in a confined visual environment with subtle vestibular input. Those two PSP and PD patients who were wheelchair-dependent were wheeled throughout the whole procedure by an experimenter instead of actively walking.

The objective of the real-life measurement was to provide a naturalistic set of behaviors, while differences between real-life conditions were not at the focus of the current study. Consequently, all data of real-life measurement were pooled per participant. The real-life measurement lasted less than 10 min per participant.

DATA ANALYSIS AND STATISTICAL EVALUATION

Eye movements

Raw eye-position data were processed offline using MATLAB (Matlab 7.10, The MathWorks, Natick, MA), which was also used for statistical analysis. We calculated eye velocity by differentiation of the horizontal and vertical eye position (**Figure 1B**). Absolute speed was then calculated as the square root of the sum of the squared horizontal and squared vertical velocity componentss.

All phases faster than 60°/s and lasting longer than 10 ms are referred to as "saccades," irrespective of whether they were actual saccades or fast phases of reflexive movements (**Figure 1B**). This threshold is higher than those typically used in laboratory settings, as signals obtained during real-life measurements contain rich eye movement dynamics and are typically noisier than under constrained settings. The conservative choice is, however, consistent with previous research on eye movements in PSP patients: for example, judging from the figures in Pinkhardt et al. (2008), their patients had their 5% percentile of peak saccade velocities around or above 60°/s, meaning that we can still expect to include about 95% of actual saccades with our comparably conservative criterion. Since this criterion could also be employed in practice, it will not affect any conclusion on the discriminability of patient groups. Nonetheless, for the general questions pertaining to eye movement disturbances in PSP and PD, the fact that any threshold must remain arbitrary motivates to add an analysis that does not classify eye movements in saccade/non-saccade, but uses the unclassified (i.e., raw) eye movement data (see below and section "Unclassified Eye Movements").

Parameters to describe saccades were their direction, peak velocity, amplitude, and duration (**Figure 1B**). Since peak velocity, saccade amplitude, and duration are typically not independent, the functional relationship of amplitude and peak velocity and of amplitude and saccade duration, the so-called main sequence (Bahill et al., 1975), was also considered for real-life data: we fitted the relation with a power function of the form velocity $= a \times \text{amplitude}^b$ or duration $= a \times \text{amplitude}^b$, respectively (cf. Garbutt et al., 2003), and considered only the fit parameters a and b further. Since reliable fits of main sequences require substantial amounts of data, this analysis was only performed for the real-life measurements.

To test whether there is an abundance of one saccade direction in a group, we coarsely classified saccades into equally spaced 45° wedges: horizontal (±22.5° from the horizontal), vertical (±22.5° from the vertical), and oblique (the remaining $4 \times 45° = 180°$).

For analysis of raw ("unclassified") eye data (i.e., all data irrespective of whether defined as saccade or not),

two-dimensional histograms were used. Each bin of the histograms used for analysis corresponds to a velocity interval of 15°/s in each direction (horizontal and vertical); the central bin ranges from −7.5°/s to +7.5°/s in each direction. The number of samples in each bin is color-coded.

Head movements

Head movements were computed from the video of the head-fixed camera at 25 Hz. To obtain head position, the same stationary point of the environment was marked in each video-frame. From this point's position in the camera's field of view relative head orientation in the world was computed. Head velocity was obtained by differentiation of this signal, and was thus independent of this choice of origin. All quantitative analysis was therefore based on velocities. Unlike for eye movements and due to the low spatial and temporal resolution (section "Eye and Head Movement Recordings" top), we could not classify head movements in distinct classes (e.g., fast/slow) with the data at hand. Therefore, all analysis was based on overall velocity distributions for each individual.

Statistical analysis

Data are presented as mean ± standard deviation. Statistical evaluation used non-parametric tests for raw eye data, such as amplitude and peak velocity of each saccade (Kruskal–Wallis when three groups were compared and Mann–Whitney-*U*-Test for two groups). To compare these parameters in an exploratory manner across participants, the individual distributions are described by their medians as robust measure (since the distributions are either leptokurtic or prone to outliers). Since these medians can be assumed to follow a normal distribution across participants, the group effects were analyzed by parametric tests; that is, ANOVAs for three group comparisons and two-tailed *t*-tests for two-group comparisons and *post-hoc* tests.

Signal-detection-theory measures

For assessing the performance of the classifiers between PSP and PD patients, we performed signal-detection analysis by computing the Receiver-Operating-Characteristic (ROC). The ROC is quantified by its area under the curve (AUC), the cut-off point for maximal specificity and sensitivity, and the corresponding values of specificity and sensitivity. Values are reported such that all values of patients classified as PSP patients are strictly smaller than this cut-off value.

RESULTS

PARTICIPANT CHARACTERISTICS

We investigated 10 PSP patients (6 probable, 4 possible), 11 PD patients and 10 HCs (**Table 1**). All patients were under treatment in the University Hospital in Marburg. There were no significant differences regarding age, disease duration, and gender between the groups. For all patients Hoehn and Yahr stage was assessed in off-state and, as expected, the stages differed significantly between PSP and PD patients (**Table 1**).

Eye velocities and relative eye positions (e.g., saccade amplitudes) require only minimal subject–specific adjustment

Table 1 | Clinical characteristics of the participants in this study: overview.

	PSP	PD	HC
N	10	11	10
Age (years)	65.9 ± 4.6	65.5 ± 12.7	68.3 ± 9.1
Gender (F/M)	3/7	3/8	6/4
DD (years)	3.9 ± 2.7	6.2 ± 4.7	–
H&Y	3.9 ± 0.4	2.5 ± 0.4	
Wheelchair	2/10	2/11	0/10
Real-life measurement time	304.3 ± 114.4 s	242.2 ± 78.5 s	202.8 ± 35.3 s

Details

Patient ID/gender/age [years]	Onset	Exam. date	H&Y	Medication
PSP01/F/67	2004	08/2010	4	Levodopa
PSP02/M/70	2008	08/2010	3	Amantadine
PSP03/F/63	2007	08/2010	4	Levodopa, Amantadine
PSP04/M/70	2007	08/2010	4	Levodopa, Amantadine, Piribedil
PSP05/F/65	2007	08/2010	3	Amantadine, Rotigotine
PSP06/M/67	2000	08/2010	4	Levodopa
PSP07/M/62	2008	02/2011	4	Levodopa
PSP08/M/74	2005	05/2011	4	Levodopa, Amantadine
PSP09/M/59	2010	10/2011	3	Levodopa
PSP10/M/62	2009	11/2011	3	Levodopa
PD01/M/61	2007	09/2010	2	Rotigotine
PD02/M/75	1995	09/2010	3	Levodopa
PD03/M/75	2007	02/2011	1	Ropinirole
PD04/M/64	2000	07/2011	3	Levodopa, Amantadine, Pramipexole, Rasagiline
PD05/M/67	2007	07/2011	1	Levodopa, Ropinirole, Rasagiline
PD06/M/51	2010	09/2011	2	Levodopa, Rasagiline
PD07/F/62	2007	10/2011	3	Levodopa, Rasagiline, Piribedil
PD08/M/38	2010	10/2011	2	Pramipexole
PD09/M/78	2007	12/2011	3	Levodopa
PD10/F/82	2001	12/2011	3	Levodopa, Amantadine, Ropinirole
PD11/F/68	2000	12/2011	3	Levodopa, Amantadine, Pramipexole
HC01/F/58		08/2010		
HC02/M/71		08/2010		

(Continued)

Table 1 | Continued

Patient ID/gender/age [years]	Onset	Exam. date	H&Y	Medication
HC03/F/53		02/2011		
HC04/F/63		02/2011		
HC05/M/73		03/2011		
HC06/F/64		03/2011		
HC07/F/69		03/2011		
HC08/F/74		09/2011		
HC09/M/85		12/2011		
HC10/M/73		12/2011		

PSP, progressive supranuclear palsy; PD, Parkinson's disease; HC, healthy controls; DD, disease duration; H&Y, Hoehn and Yahr Stage. H&Y stage is significantly different between PD and PSP [$t_{(19)} = 4.12$, $p < 0.001$]; real-life measurement duration differs significantly between PSP and HC ($p = 0.02$ post-hoc test); all other comparisons do not show a significant difference ($p > 0.05$).

and could thus be measured accurately in all participants. However, individual-specific calibration of absolute eye-position failed in eight PSP and two PD patients as a consequence of their inability to steadily fixate instructed targets over a 2-s integration window. Interestingly, this inability did not primarily result from square-wave jerks, which were robustly observed only in 1 out of the 10 PSP patients under our experimental conditions. As a consequence of the calibration failures for absolute position, all quantitative analysis hereafter is based on relative eye-position and velocities only.

SACCADES

Fixation protocol

All participants performed a standard fixation protocol, as described in the "Materials and Methods" section, which was also used for individual calibration refinement. Irrespective of whether this absolute-position calibration was successful or not, these measurements provided a sufficient number of visually-guided saccades to analyze differences between PSP patients and PD patients or HCs (**Figure 2**).

Averaged median saccadic peak velocity was $135.1 \pm 43.8°$/s for PSP, $220.1 \pm 31.5°$/s for PD patients and $233.0 \pm 44.4°$/s for HCs. A One-Way ANOVA revealed a significant main effect [$F_{(2,\ 28)} = 17.81$, $p < 0.001$, **Figure 2B**] and *post-hoc t*-tests showed that PSP patients generated saccades with significantly slower median peak velocity than PD patients [$t_{(19)} = 5.14$, $p < 0.001$] and HCs [$t_{(18)} = 4.96$, $p < 0.001$]. There were also significant differences in the vertical components of saccade peak velocity. Averaged vertical saccade peak velocity was $54.9 \pm 28.0°$/s for PSP patients, $158.5 \pm 47.9°$/s for PD patients and $151.1 \pm 60.3°$/s for HCs [$F_{(2,\ 28)} = 14.53$, $p < 0.001$; PSP-PD: $t_{(19)} = 5.83$, $p < 0.001$; PSP-HC: $t_{(18)} = 4.51$, $p < 0.001$, **Figure 2C**].

Saccade amplitudes also differed significantly between groups [$F_{(2,\ 28)} = 18.26$, $p < 0.001$, PSP-PD: $t_{(19)} = 4.26$, $p < 0.001$, PSP-HC: $t_{(18)} = 6.60$, $p < 0.001$, **Figure 2B**]. Averaged median amplitudes were $1.88 \pm 0.72°$ for PSP patients, $4.16 \pm 1.53°$ for PD patients and $5.42 \pm 1.53°$ for HCs. Vertical saccade amplitude was $0.52 \pm 0.37°$ for PSP patients, $2.89 \pm 1.62°$ for PD patients and $3.03 \pm 2.16°$ for HCs and thus also differed significantly [$F_{(2,\ 28)} = 7.76$, $p = 0.002$; PSP-PD: $t_{(19)} = 4.37$, $p < 0.001$; PSP-HC: $t_{(18)} = 3.57$, $p = 0.002$, **Figure 2C**].

We did not find significant main effects for the horizontal components of peak velocity [$F_{(2,\ 28)} = 2.12$, $p = 0.14$, ANOVA; **Figure 2D**] and amplitude [$F_{(2,\ 28)} = 1.69$, $p = 0.20$, **Figure 2D**].

The ROC comparing saccade peak velocity of PSP and PD patients showed an AUC of 0.95. Specificity was 11/11 and sensitivity was 9/10 for a cut-off value of 189.8°/s (i.e., all patients having slower peak velocities than this value were classified as PSP) patients. For the comparison of vertical saccade peak velocities, the AUC was 1 and for the cut-off value 111.7°/s, specificity was 11/11 and sensitivity was 10/10. The AUC for the comparison of saccade amplitude was 0.97 with a specificity of 11/11 and a sensitivity of 9/10 for a cut-off value of 2.79°. For the vertical component, AUC was 0.99 and the ROC analysis showed a specificity of 10/11 and a sensitivity of 10/10 for the cut-off value 1.68°.

For completeness, we also analyzed saccade duration in all groups. We found a significant main effect between groups [PSP: 19.6 ± 7.2 ms, PD: 26.2 ± 6.3 ms, HC: 32.7 ± 6.5 ms, $F_{(2,\ 28)} = 9.6$, $p < 0.001$, see **Figures 2E,F**]. *Post-hoc t*-test revealed significant differences between all groups [PSP-PD: $t_{(19)} = 2.25$, $p = 0.037$; PSP-HC: $t_{(18)} = 4.27$, $p < 0.001$; PD-HC: $t_{(19)} = 2.30$, $p = 0.033$]. Sensitivity was 7/10 and specificity was 9/11 for the cut-off value 21.6 ms, the AUC was 0.77. These values are much lower than for amplitude and peak velocity and thus less informative where differential diagnosis is concerned. Hence, we hereafter focus most analysis on peak velocity and amplitude.

Real-life

Since the eye movement impairment in PSP was evident during the fixation protocol, we next analyzed their relevance for real-life situations. Hence, we measured the spontaneous ocular motor behavior in a real-life, minimally restrained scenario, comprising self-paced walking in a corridor, tracking of a stationary target, and taking an elevator. Self-paced walking implies speed differences between participants. ANOVA revealed a significant main effect for differences in real-life measurement duration [$F_{(2,\ 28)} = 3.85$, $p = 0.03$, **Table 1**]; the difference was not significant between PSP and PD patients, but for HCs the measurement lasted significantly shorter than for PSP patients [$t_{(18)} = 2.68$, $p = 0.02$]. Aggregating over the whole real-life measurement, we assessed the same parameters as during the fixation protocol (**Figure 3**).

All groups had the same fraction of vertical [PSP: 24.1% ± 15.4%, PD: 28.9% ± 10.4%, HC: 31.7% ± 7.1%, $F_{(2,\ 28)} = 1.14$, $p = 0.33$], horizontal [PSP: 21.7% ± 9.0%, PD: 18.5% ± 8.1%, HC: 18.3% ± 5.9%, $F_{(2,\ 28)} = 0.64$, $p = 0.53$] and oblique [PSP: 54.3% ± 10.1%, PD: 52.6% ± 6.1%, HC: 50.0% ± 3.7%, $F_{(2,\ 28)} = 0.92$, $p = 0.41$] saccades.

The medians of saccade peak velocity differed significantly between the groups [$F_{(2,\ 28)} = 5.47$, $p = 0.01$, **Figure 3B**]. PSP

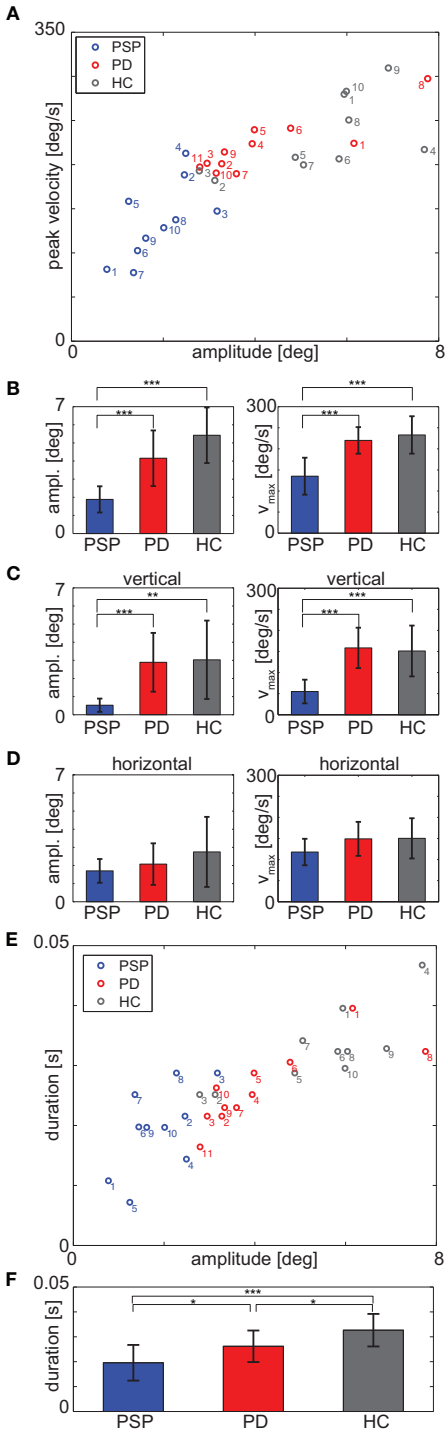

FIGURE 2 | Continued

FIGURE 2 | (A) Medians of saccade peak velocity and amplitude for each participant during the fixation protocol. **(B)** Mean over participants of median amplitude (left panel) and median peak velocity (right panel) for each group. **(C)** Vertical component and **(D)** horizontal component of the data of panel **(B)**; **(E)** Medians of saccade duration and amplitude for each participant during fixation protocol; note that the duration is discretized due to sampling frequency **(F)**. Mean over participants of median duration. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

patients' averaged median saccade peak velocity was 131.1 ± 29.0°/s and thus slower than those of PD patients [163.1 ± 25.8°/s; $t_{(19)} = 2.68$, $p = 0.002$] and HCs [160.2 ± 15.4°/s; $t_{(18)} = 2.80$, $p = 0.01$]. The vertical component of saccade peak velocity (PSP: 71.9 ± 15.5°/s, PD: 89.6 ± 11.5°/s, HC: 89.5 ± 9.6°/s) also differed significantly [$F_{(2,\ 28)} = 6.88$, $p = 0.004$, PSP-PD: $t_{(19)} = 3.00$, $p = 0.007$; PSP-HC: $t_{(18)} = 3.05$, $p = 0.007$, **Figure 3C**], whereas there was no significant difference between means of the horizontal component of peak velocity [$F_{(2,\ 28)} = 1.66$, $p = 0.21$, **Figure 3D**] between groups.

ANOVA did not reveal a significant main effect for saccade amplitude [$F_{(2,\ 28)} = 2.55$, $p = 0.10$, **Figure 3B**], but the vertical component of saccade amplitude differed significantly [$F_{(2,\ 28)} = 3.46$, $p = 0.045$, **Figure 3C**]; *post-hoc t*-tests revealed that PSP patients' vertical component of saccade amplitude was significantly shorter (0.79 ± 0.36°) than PD patients' [1.12 ± 0.33°; $t_{(19)} = 2.12$, $p = 0.047$] and HCs' [1.06 ± 0.13°; $t_{(18)} = 2.16$, $p = 0.04$]. There was no significant difference between medians of the horizontal components of amplitudes [$F_{(2,\ 28)} = 0.25$, $p = 0.78$, **Figure 3D**].

The AUC was 0.84 for peak velocity with a sensitivity of 8/10 and a specificity of 9/11 for the cut-off value 139.9°/s. For vertical peak velocity, the AUC was 0.82 and for a cut-off value of 83.2°/s sensitivity was 7/10 and specificity was 8/11. For analysis of saccade amplitudes, the AUC was 0.80 with a sensitivity of 8/10 and a specificity of 8/11 for a cut-off value of 1.85°. The AUC for comparison of vertical components was 0.75 with a sensitivity of 6/10 and a specificity of 11/11 for the cut-off value 0.69°.

Differences in medians of saccade duration were not significantly different between groups [PSP: 25.5 ± 3.7 ms, PD: 27.6 ± 4.0 ms, HC: 25.6 ± 2.3 ms, $F_{(2,\ 28)} = 1.31$, $p = 0.29$; see **Figures 3E,F**].

Correlation between fixation protocol and real-life

Median of peak velocity and its vertical component in the fixation protocol and during real-life measurement correlated significantly ($N = 31$, $r = 0.39$, $p = 0.03$; vertical: $r = 0.50$, $p = 0.004$). Thus, the data collected during the fixation protocol not only differentiated between PSP and PD patients, but also in part predicted real-life performance.

Main-sequence analysis

Peak velocity and duration were plotted as a function of amplitude for each saccade of every participant. We fitted this main sequence with a power function (**Figure 4A**) and compared the fit parameters between groups. There were no significant differences between groups with respect to the value of fit parameters *a*

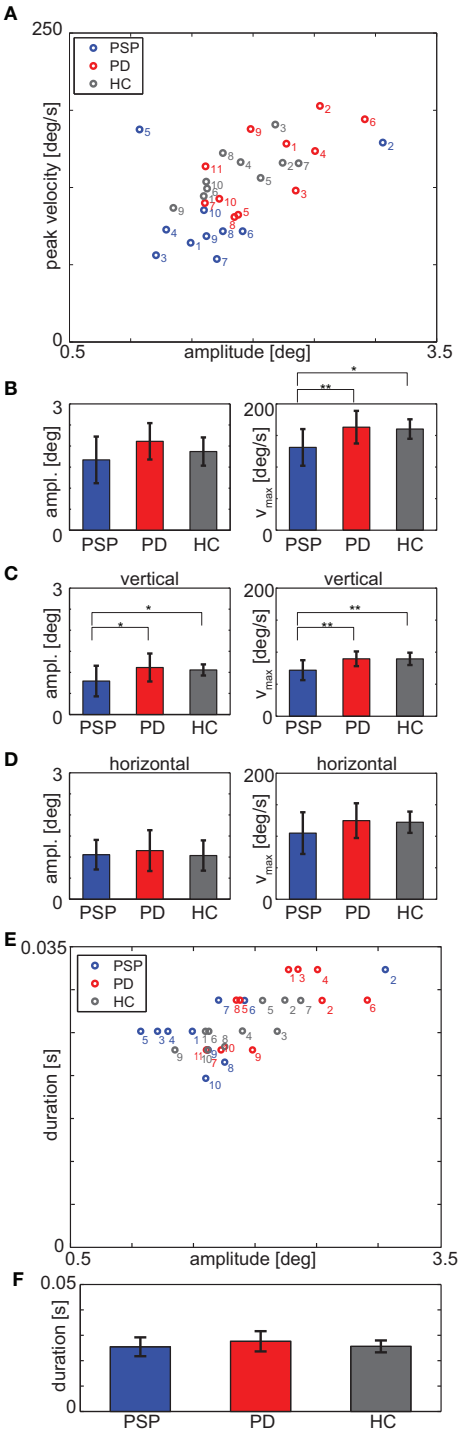

FIGURE 3 | Continued

FIGURE 3 | (A) Medians of saccade peak velocity and amplitude for each participant during real-life measurement. **(B)** Mean over participants of median amplitude (left panel) and median peak velocity (right panel) for each group. **(C)** Vertical component and **(D)** horizontal component of the data of panel **(B)**. **(E)** Medians of saccade duration and amplitude for each participant during real-life measurement; note that the duration is discretized due to sampling frequency **(F)**. Mean over participants of median duration. $^{*}p < 0.05$; $^{**}p < 0.01$.

$[F_{(2,\ 28)} = 1.69, p = 0.20$, **Figure 4B**$]$ and b $[F_{(2,\ 28)} = 1.38, p = 0.27$, **Figure 4C**$]$. There were also no differences between groups in the vertical component of saccades [value of a: $F_{(2,\ 28)} = 2.54, p = 0.097$, **Figure 4D**; value of b: $F_{(2,\ 28)} = 1.08, p = 0.35$, **Figure 4E**] and in the value of the fit parameter a of the functional relationship between duration and amplitude $[F_{(2,\ 28)} = 0.02, p = 0.98$, **Figure 4F**$]$. There was a significant main effect for the values of b in that case $[F_{(2,\ 28)} = 4.11, p = 0.027$, **Figure 4G**$]$ but *post-hoc* t-tests did not reveal significant differences between PSP and PD patients $[t_{(19)} = 1.77, p = 0.09]$ or PD patients and HCs $[t_{(19)} = 1.24, p = 0.23]$. The only significant difference was found between PSP patients and HCs $[t_{(18)} = 2.43, p = 0.026]$.

UNCLASSIFIED EYE MOVEMENTS

Under real-life conditions, fast eye movement phases (saccades), as analyzed above, accounted for only a small amount of the entire measurement time (PSP: $7.6 \pm 3.8\%$, PD: $11.7\% \pm 7.9\%$, HC: $10.4\% \pm 2.8\%$). To compare saccade-based analysis to all eye movements, we generated 2-dimensional velocity histograms for saccades only (**Figure 5A**) and for all eye movements ("unclassified movements," **Figure 5B**) during the entire real-life measuring time. The histograms show pooled data from all participants of each group, normalized such that each participant contributes with equal weight to the respective histograms. In the distribution of saccade peak velocities (**Figure 5A**), a preference for horizontal movements is evident in all groups, which is particularly pronounced in PSP patients, reflecting their prominent reduction in vertical peak velocity. Interestingly, this difference between groups was less evident when analyzing all eye movements (**Figure 5B**). We quantified the spread in each direction by standard deviation. When considering all unclassified eye movements, there were no significant differences among the groups [vertical: $F_{(2,\ 28)} = 1.74$, $p = 0.19$; horizontal: $F_{(2,\ 28)} = 1.86$, $p = 0.18$]. When instead considering saccades only (**Figure 5A**), a picture consistent with the analysis above (section "Real-Life") emerged: the standard deviation of saccade peak velocities yielded highly significant differences between the groups [vertical: $F_{(2,\ 28)} = 8.53$, $p = 0.001$; horizontal: $F_{(2,\ 28)} = 12.42$, $p < 0.001$]. Significant differences appeared between PSP and PD patients [vertical: $t_{(19)} = 3.38$, $p = 0.003$; horizontal: $t_{(19)} = 4.34$, $p < 0.001$] as well as between PSP patients and HCs [vertical: $t_{(18)} = 3.41$, $p = 0.003$; horizontal: $t_{(18)} = 3.75$, $p = 0.002$]. Moreover, when testing analogous measures to those that yielded significant differences and high diagnostic power between patient groups for saccades (**Figures 2** and **3**), no significant effects were found for the full, unclassified eye movement data. For example, the medians of all velocities were not significantly different between the groups

2.5 Validation of mobile eye-tracking as novel and efficient means for differentiating progressive supranuclear palsy from Parkinson's disease

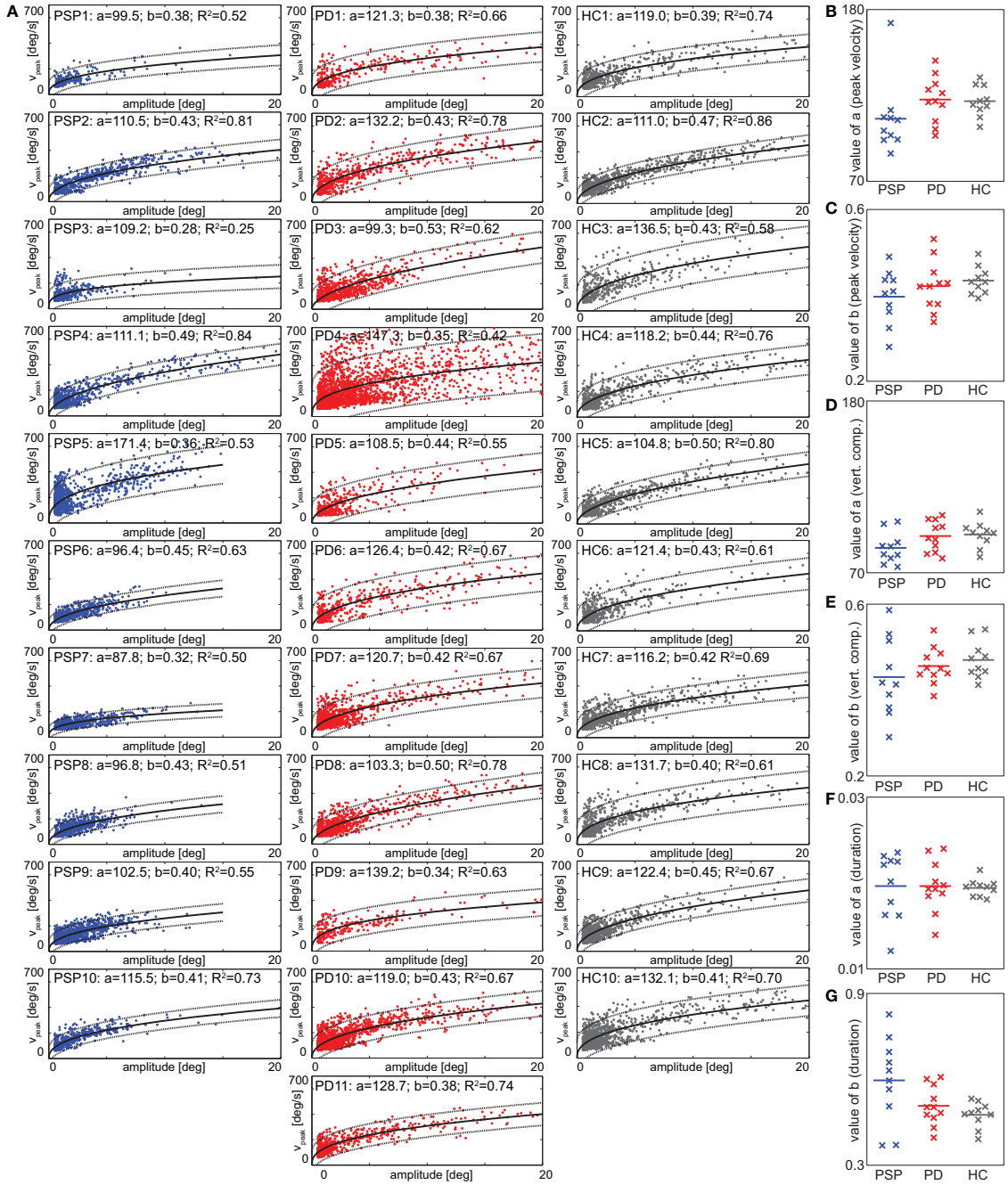

FIGURE 4 | (A) Saccade peak velocity plotted against saccade amplitude ("main sequence") for all individuals. Each data-point corresponds to one saccade (note the cutoff to the bottom and left, given by the thresholds on velocity and duration). Solid black line denotes best fitting power functions (see section "Eye Movements") in a least-squares sense, dotted lines 5 and 95% confidence intervals. Fit parameters and R^2 are given in the panel headers. **(B,C)** Main-sequence fit parameters of the functional relationship between peak velocity and amplitude. **(D,E)** Main-sequence fit parameters for the functional relationship between vertical component of peak velocity and amplitude. **(F,G)** Main-sequence fit parameters for the functional relationship between saccade duration and amplitude.

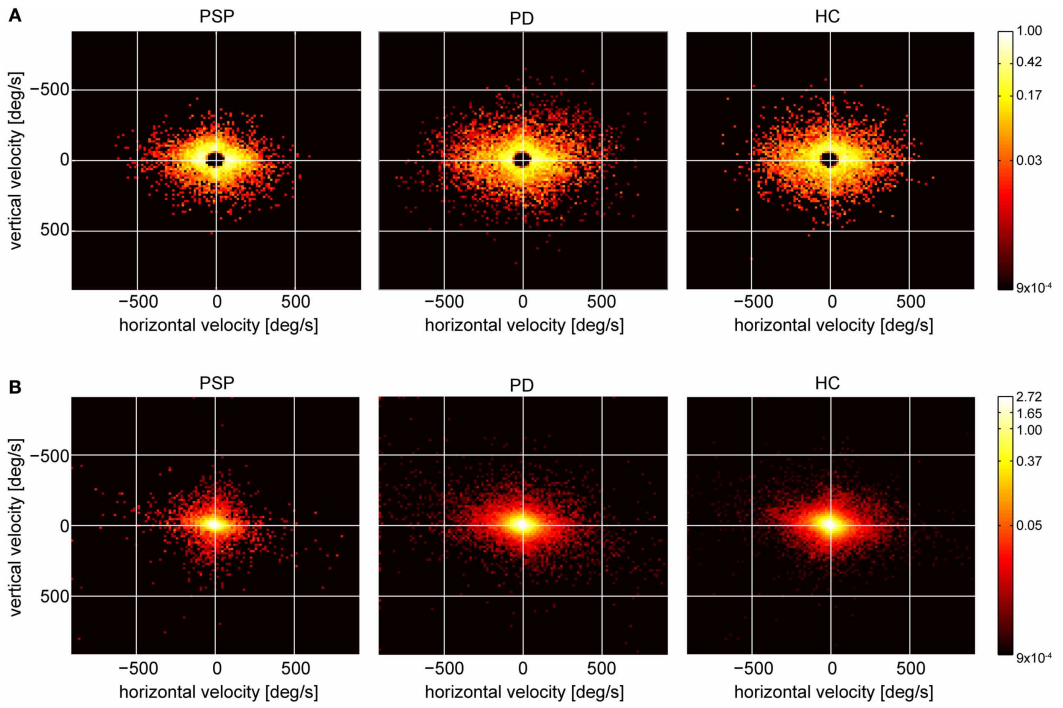

FIGURE 5 | (A) Peak velocity histograms of cumulated saccades of all participants in each group (range: −900 to 900°/s in the cardinal directions, bin size: 15°/s × 15°/s). **(B)** Velocity histograms of raw eye velocities of all cumulated data points of all participants in each group, same range, and bin size as in **(A)**.

$[F_{(2,\ 28)} = 1.01, p = 0.38]$. Notwithstanding some degree of arbitrariness in the definition of saccade thresholds, this indicates that—at least under our recording conditions—the described effects are best observed in fast movements.

HEAD MOVEMENTS

For 26 participants (9 PSP, 7 PD, and 10 HC) we successfully obtained head data during the fixation protocol, for 27 (9 PSP, 9 PD, and 9 HC) during walking along the corridor without target tracking, and for 29 (9 PSP, 10 PD, and 10 HC) while they tracked the stationary target. In the remaining participants, head orientation was not recorded or recording was unsuccessful for technical reasons. We chose to split walking the corridor into periods with tracking and without tracking for head-in-world data considered here, as we expected higher consistency with respect to the overall head movements.

During the fixation protocol, all but one participant deviated less than 2° from their average gaze orientation, 22/26 even less than 1°. Thus, head movements were small and rare, and the median head velocity was below 2°/s in all but one participant. While this implies that participants complied with the instruction to avoid head movements, it also means insufficient movements to obtain robust velocity data.

During tracking, spread (quantified as standard deviations) of head velocities was not significantly different between groups [vertical: $F_{(2,\ 26)} = 0.49$, $p = 0.62$, **Figure 6A**; horizontal: $F_{(2,\ 26)} = 0.63$, $p = 0.54$, **Figure 6B**]. During walking without tracking, the vertical spread in velocity showed no dependence on group [$F_{(2,\ 24)} = 0.51$, $p = 0.61$, **Figure 6C**], either. In contrast, horizontal spread showed a significant group dependence [$F_{(2,\ 24)} = 3.67, p = 0.04$, **Figure 6D**], indicating that the absence of an effect during tracking, where less participants contributed, was not due to a lack of power. Importantly, this group dependence resulted from a difference between PSP patients and HCs [PSP-HC: $t_{(16)} = 3.41$, $p = 0.004$], but not from a difference between patient groups [PSP-PD: $t_{(16)} = 0.01$, $p = 0.99$] or between PD patients and HCs [PD-HC: $t_{(16)} = 2.07$, $p = 0.055$]. In sum, neither head orientation nor head velocity—to the extent they could be analyzed with the present device—could offer any parameters that might serve to discriminate PSP from PD.

DISCUSSION

In the present study we used a novel, wearable eye-tracking device to assess gaze behavior in PD, PSP, and HCs. First, we demonstrate that wearable eye-tracking distinguishes PSP from PD with high sensitivity and specificity. Second, we show that these differences in gaze behavior are most prominent for saccades in a brief fixation protocol and less pronounced in activities of daily living.

The observed differences between saccadic peak velocities in the fixation protocol are highly consistent with earlier findings

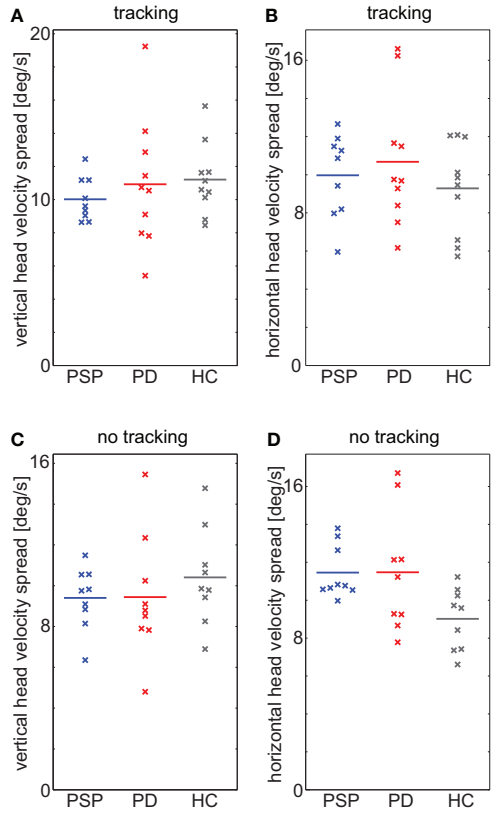

FIGURE 6 | (A,B) Spread of head velocity in **(A)** vertical and **(B)** horizontal direction during walking while tracking a stationary target. **(C,D)** Spread of head velocity in **(C)** vertical and **(D)** horizontal direction during walking without specific instruction. Horizontal displacement of datapoints within groups is for improved visibility only.

(Pinkhardt and Kassubek, 2011; Boxer et al., 2012). Similarly, the lack of evidence for a difference in peak velocities between the PD group and HCs are in line with previous data (Tanyeri et al., 1989; Pinkhardt and Kassubek, 2011). As such, our data extend earlier findings obtained using visually-guided saccades in standard laboratory setups to wearable eye-tracking, which allows efficient assessment of these parameters in less restrained conditions. Even though many sorts of eye movements are affected by PSP, we focused on saccadic peak velocity and amplitude for reasons of efficiency. Duration of saccades as conceivable alternative turned out to have less diagnostic power, despite some difference in the average. Although amplitude, peak velocity, and duration are not independent, but coupled through the "main sequence," the functional fit does not provide any additional diagnostic power in real-life data, and requires more data than available from the 20-s fixation protocol, such that amplitude and peak velocity remain as the main diagnostic markers for this rapid assessment. Still, if these two parameters should turn out to be insufficient for differential diagnosis in a patients with clinically uncertain diagnosis, other eye movements like vergence and the linear vestibuloocular reflex can also be measured with the EyeSeeCam.

The comparison between raw data and data filtered for saccades allows three main conclusions. First, it stresses the specifically prominent impairment of the saccade system for PSP patients as compared to other eye movement systems (Chen et al., 2010). Second, it underlines the importance of objective measurement devices to reliably detect potentially subtle eye movement-related disease markers (Bartl et al., 2009). Finally, the comparably mild differences in overall gaze orienting behavior might point to a strategy how the specific deficits may be compensated for and thus offers a promising path for carefully quantifiable therapeutic intervention (Zampieri and Di Fabio, 2008).

The reduced differences in gaze behavior during activities of daily living indicate that patients at least in part compensate for their ocular motor deficits. Analysis of head movements, however, suggests substantial inter-individual differences, indicating that compensation strategies are largely idiosyncratic. Predicting such compensation behaviors and relating them to other parameters, such as disease progression, will be an interesting issue for further research in larger, heterogeneous PSP cohorts. In a longitudinal study, the precise quantification of compensatory behavior might then also aid the efficient monitoring of treatment success. For differential diagnosis, the free exploration paradigm is clearly less valuable, demonstrating the importance of a flexible, but at the same time standardized fixation protocol for clinical use. Nonetheless, the free exploration data may yield important information on compensation mechanisms and the consequences of the disease on everyday life.

In contrast to eye movements, the parameters considered for head movements did not allow a significant dissociation between patient groups under any of the tested tasks. This could be due to the low spatial and temporal resolution of the head movement measurements as compared to eye movement measurements. It is conceivable that with an improved measurement device for head movements, with different instructions or tasks, or when effects on eye-head coordination are measured with sufficient spatial and temporal accuracy and precision, head movements might eventually become useful and could augment a PSP/PD discrimination system. However, with the present technology and based on the tasks used in the present study, eye velocity and amplitude during the fixation protocol present a most promising candidate for dissociating PSP from PD also in subclinical populations.

This study is to be regarded as a first step toward establishing a new method as a diagnostic tool. Prospective studies measuring eye movements of still unclassified patients are needed to prove that subclinical oculomotor disturbances can be detected prior to the establishment of the clinical diagnosis. Also, square wave jerks which are characteristic of PSP patients could only be detected in one PSP patient, even by careful visual inspection of all eye movement traces. While beyond the scope of the present study, the question as to whether their absence from the measured data is a technical limitation or a true effect of the

population and condition at hand remains an important issue for future research.

Importantly for a possible application in diagnosis and treatment monitoring, the usage of the wearable eye-tracking device is efficient, requiring less than 20-s for the fixation protocol and virtually no device-specific training. While wearable eye-tracking has recently been suggested as tool in a variety of ocular motor and vestibular conditions (Hayhoe and Ballard, 2005; Schumann et al., 2008), the present study demonstrates that wearable eye-tracking also lends itself for efficient clinical use in the context of more complex syndromes, such as typical and atypical Parkinsonism. Whether or not wearable eye-tracking will allow diagnosis beyond the current gold standard obviously can only be established in a long-term longitudinal prospective study, which will apply the criteria found herein already early during disease, when current clinical criteria are not yet clear cut.

ACKNOWLEDGMENTS

Financial support by the German Research Foundation (DFG) through grant EI 852/1 (Wolfgang Einhäuser), the Swiss National Science Foundation (SNF) and DFG through grant 135687/EI 852/3 (Wolfgang Einhäuser), and the DFG through grant HO2402/6-1 (Günter U. Höglinger) is gratefully acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http:// www.frontiersin.org/Behavioral_Neuroscience/10.3389/fnbeh.2012.00088/abstract

Movie 1 | Example movies of two participants, PD07 and PSP09, showing a part of the real-life measurement. Histograms picture eye velocity (left panel, range: −500 to 500°/s in the cardinal directions, bin size for this movie: 5°/s × 5°/s) and head velocity (right panel, range: −60 to 60°/s in the cardinal directions, bin size: 3°/s × 3°/s).

REFERENCES

Bahill, A. T., Clark, M. R., and Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Math. Biosci.* 24, 191–204.

Bartl, K., Lehnen, N., Kohlbecher, S., and Schneider, E. (2009). Head impulse testing using video-oculography. *Ann. N.Y. Acad. Sci.* 1164, 331–333.

Boxer, A. L., Garbutt, S., Seeley, W. W., Jafari, A., Heuer, H. W., Mirsky, J., et al. (2012). Saccade abnormalities in autopsy-confirmed frontotemporal lobar degeneration and Alzheimer disease. *Arch. Neurol.* 69, 509–517.

Brandt, T., Glasauer, S., and Schneider, E. (2006). A third eye for the surgeon. *J. Neurol. Neurosurg. Psychiatry* 77, 278.

Burn, D. J., and Lees, A. J. (2002). Progressive supranuclear palsy: where are we now? *Lancet Neurol.* 1, 359–369.

Chen, A. L., Riley, D. E., King, S. A., Joshi, A. C., Serra, A., Liao, K., et al. (2010). The disturbance of gaze in progressive supranuclear palsy: implications for pathogenesis. *Front. Neur.* 1:147. doi: 10.3389/fneur.2010.00147

Garbutt, S., Harwood, M. R., Kumar, A. N., Han, Y. H., and Leigh, R. J. (2003). Evaluating small eye movements in patients with saccadic palsies. *Ann. N.Y. Acad. Sci.* 1004, 337–346.

Gibb, W. R., and Lees, A. J. (1988). The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 51, 745–752.

Golbe, L. I., and Ohman-Strickland, P. A. (2007). A clinical rating scale for progressive supranuclear palsy. *Brain* 130, 1552–1565.

Hayhoe, M., and Ballard, D. (2005). Eye movements in natural behavior. *Trends Cogn. Sci.* 9, 188–194.

Litvan, I., Agid, Y., Calne, D., Campbell, G., Dubois, B., Duvoisin, R. C., et al. (1996). Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome). *Neurology* 47, 1–9.

Otero-Millan, J., Serra, A., Leigh, R. J., Troncoso, X. G., Macknik, S. L., and Martinez-Conde, S. (2011). Distinctive features of saccadic intrusions and microsaccades in progressive supranuclear palsy. *J. Neurosci.* 31, 4379–4387.

Pinkhardt, E. H., Jurgens, R., Becker, W., Valdarno, F., Ludolph, A. C., and Kassubek, J. (2008). Differential diagnostic value of eye movement recording in PSP-parkinsonism, Richardson's syndrome, and idiopathic Parkinson's disease. *J. Neurol.* 255, 1916–1925.

Pinkhardt, E. H., and Kassubek, J. (2011). Ocular motor abnormalities in Parkinsonian syndromes. *Parkinsonism Relat. Disord.* 17, 223–230.

Schneider, E., Bartl, K., Bardins, S., Dera, T., Boning, G., and Brandt, T. (2006). Documentation and teaching of surgery with an eye movement driven head-mounted camera: see what the surgeon sees and does. *Stud. Health Technol. Inform.* 119, 486–490.

Schneider, E., Villgrattner, T., Vockeroth, J., Bartl, K., Kohlbecher, S., Bardins, S., et al. (2009). Eyeseecam: an eye movement-driven head camera for the examination of natural visual exploration. *Ann. N.Y. Acad. Sci.* 1164, 461–467.

Schumann, F., Einhäuser, W., Vockeroth, J., Bartl, K., Schneider, E., and König, P. (2008). Salient features in gaze-aligned recordings of human visual input during free exploration of natural environments. *J. Vis.* 8, 12.1–12.17.

Tanyeri, S., Lueck, C. J., Crawford, T. J., and Kennard, C. (1989). Vertical and horizontal saccadic eye movements in parkinson's disease. *Neuroophthalmology* 9, 165–177.

Zampieri, C., and Di Fabio, R. P. (2008). Balance and eye movement training to improve gait in people with progressive supranuclear palsy: quasi-randomized clinical trial. *Phys. Ther.* 88, 1460–1473.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 October 2012; accepted: 27 November 2012; published online: 13 December 2012.

Citation: Marx S, Respondek G, Stamelou M, Dowiasch S, Stoll J, Bremmer F, Oertel WH, Höglinger GU and Einhäuser W (2012) Validation of mobile eye-tracking as novel and efficient means for differentiating progressive supranuclear palsy from Parkinson's disease. Front. Behav. Neurosci. 6:88. doi: 10.3389/fnbeh.2012.00088

General Discussion

In a series of five studies, perceptual phenomena and mechanisms have been studied using complex - ambiguous or natural - stimuli. Furthermore, the application of eye-movement measurements in the real world has been demonstrated. In a first experiment, we showed that reward exerts an enhancing influence on perception in binocular rivalry in addition to an attentional effect. In a second study, it was demonstrated that ambiguous vision can be modeled using a winner-take-all (WTA) model frequently used to model attentional phenomena, emphasizing the link between attention and rivalry. Third, influences of color on rapid animal detection, recognition and the attentional blink (AB) were tested. In detection, a benefit of color over grayscale and color inverted pictures was found for most stimulus onset asynchronies (SOAs), but not for categorization or the AB. In the fourth study, it was shown that eye-movement parameters are impaired in older as compared to younger adults. In the final study, mobile eye tracking was brought to application by using it for differential diagnosis of two Parkinsonian syndromes. In this study we also showed that visual impairments are more pronounced in a standard fixation task than in the real-life measurement indicating compensatory mechanisms employed by the patients in real-life situations.

Competition and priority control

In studies 1 and 2, competition and its manifestation in attention and perception was investigated. The effects of attention on binocular rivalry were studied intensely in the past decade (for review see Paffen and Alais, 2011), but the effect of value on perception in rivalry was clearly demonstrated for the first time in this study. It is important to note that rivalry is not an artificial situation produced for studying visual perception in a very specific behavior. In rivalry, the inherent ambiguity of perception due to the fragmentary information we are provided with by our sensory system, is brought to the extreme. Thus it represents the "perception-as-inference" notion of perception. In real life the inference process rarely occurs between two equally likely interpretations and thus the phenomenon of rivalry is rarely experienced (for review see Arnold, 2011). However, mechanisms that are involved in everyday perception are, at least in part, also engaged in processing rival stimuli, making it a good tool to study the competitive mechanisms of perception. In our studies, two different competitive mechanisms, and how priority control can be exerted in these phenomena, were investigated: binocular rivalry,

i.e. visual competition between representations of the different stimuli presented to each eye (study 1 and 2), and attention as competition between stimuli in the visual field which can be biased towards the attended one (study 1).

We saw a strong attentional modulation of binocular rivalry in our first experiment in study 1, which contradicts some earlier studies that showed a much less pronounced attentional influence on binocular rivalry as compared to perceptual rivalry (Meng & Tong, 2004; van Ee et al., 2005). Even though we did not directly compare our binocular rivalry situation to perceptual rivalry, the large and significant influence of attention on binocular rivalry contrasts the findings of these earlier studies. They found that the influence of attention was marginally significant and differences in dominance durations between attended and unattended stimuli were not much pronounced in binocular rivalry. A potential reason for these conflicting findings is that in study 1, we used drifting gratings known to evoke eye movements, that could in turn stabilize perception in binocular rivalry. In an earlier experiment, a positive temporal correlation between saccades and alternations in perception was found (van Dam & van Ee, 2006), implying a destabilizing effect of saccades on perception. However, these were stationary stimuli so eye movements were not evoked by the stimulus itself and thus were not directly linked to the perceived interpretation itself as was the case in study 1. Hence, in our study, eye movements were percept-dependent and could thus lead to a stabilizing effect in analogy to other percept-dependent body movements that proved to stabilize perception in rivalry (Beets et al., 2010). Thus, percept-dependent eye movements could facilitate priority control in rivalry.

While selective attention and rivalry are both based on competitive mechanisms, the degree to which this competition can be resolved differs. In the framework of biased competition, competition in selective attention can be entirely resolved by setting biases, and thus priority control can be exerted. In rivalry, this is only possible temporarily for short times (the dominance durations). Even though there can be a bias towards one percept induced by attention or value in rivalry, as measured in study 1, competition is never completely resolved and the other percept will be perceived at some point after a gradual transition (Naber, Frässle, & Einhäuser, 2011). Hence, priority control can only be exerted partially in rivalry. This crucial difference is only eliminated or at least strongly reduced when the stimulus is periodically removed, which leads to stabilization of the percept (Orbach et al., 1963; Leopold et al., 2002). Since, for longer blanking periods, the percept seen before the stimulus removal is equal to the percept after the removal, memory likely plays a role in this phenomenon. This was also indicated by study 2, where we implicitly implemented memory in the WTA circuit and only this led to a good performance of the model in simulating the blanking task, whereas the memory-less circuits failed in reproducing blanking behavior. Hence, memory seems to be the factor reducing or even overcoming this limitation in resolving competition in rivalry. This seems to contradict the view that rivalry is a memory-less process with successive dominance durations being independent and

the timing of switches being unpredictable (Fox & Herrmann, 1967; Blake, Fox, & McIntyre, 1971; Levelt, 1967). However, there is, in addition to the findings on blanking, growing evidence that a correlation between successive perceptual states exists (van Ee, 2009; Naber, Gruenhage, & Einhäuser, 2010). Additionally, there are several objective physiological measures that are related to perception in rivalry such as pupil size, eye blinks, saccades and eye position (Einhäuser, Martin, & König, 2004; Einhäuser, Stout, Koch, & Carter, 2008; Naber et al., 2011; van Dam & van Ee, 2006) potentially serving as predictors of dominance durations and switch times. This renders it likely that perceptual history and memory in general do play a role in rivalry by decreasing competition and thus facilitating priority control.

The conceptual similarities between rivalry and attention also became apparent in the second study, where we used an attentional model to simulate rivalry behavior and succeeded in covering crucial rivalry characteristics. Since the model is physiologically plausible and succeeded in modeling rivalry as well as attentional behavior, it supports our view of a common competitive mechanism underlying attention, rivalry, and thus perception itself. The presented model is based on bottom-up information that consists of the noisy input applied to each input unit. By recurrently connecting two WTA circuits, this noisy input led to alternations in the perceptual circuit and correctly replicated experimental findings on the effect of stimulus strength on dominance durations and alternation rates. Most strikingly, the perceptually stabilizing effect of stimulus removal could be captured by our model making it one of the most powerful models on rivalry. Further research could test to what degree the findings on the top-down effects of attention and value can be implemented in the network. Betz, Kietzmann, Wilming, and König (2010) investigated how top-down task-dependent information interacts with bottom-up processing of stimulus-related information in a study using screenshots of webpages as stimuli and different instructions and tasks while recording eye movements. They proposed that the interaction of high-level task information and bottom-up stimulus-related information on overt visual attention could be due to one of two mechanisms: Either a weak top-down effect where top-down influences are implemented by changing feature weights in the bottom-up system or, alternatively, the strong top-down effect where the top-down influence is independent of the bottom-up processing of the stimulus. The authors tested their hypotheses using a linear bottom-up model based on image features. The findings show that the weak top-down hypothesis cannot entirely account for the task-dependent changes in viewing behavior of scenes, so there must be a direct and independent influence of top-down processes on overt visual attention, as proposed by the strong top-down hypothesis. Based on this finding, we could extend our model to also implement top-down effects by adding additional external input to the perception units according to attentional demands or reward influences.

Taken together, our findings from study 1 and 2 imply that competition - with or without priority control - is a fundamental principle of perception linking the

phenomena of rivalry, attention and memory. Furthermore, our findings imply that attention and/or memory can help resolving competition.

Attention and valuation

The first study showed that value has an effect on perception. This effect was indistinguishable from the attentional one when they are studied in isolation, but when both effects were studied in the same situation, a clear effect of reward on top of the attentional effect was demonstrated. Since it is known that reward has an impact on attention (see introduction), it can be argued that reward acts through attentional mechanisms. This would be reflected by established measures of attention like reaction time and performance. In fact, study 1 showed a tendency towards lower reaction times and higher performance for the rewarded stimulus, supporting, or at least not invalidating, this notion. However, the task engaged attentional resources quite efficiently which renders it unlikely that there were many attentional resources available to increase perceptual dominance as dramatically as reward did in study 1. It was shown that reward can increase visual salience of the rewarded stimulus (Hickey et al., 2010), which in turn makes it more likely that attention will be employed there. Hence, the argument that reward acts through attentional mechanisms could be turned around and even makes it plausible, also in our paradigm, to view attention as a reward-related mechanism (Hickey et al., 2010). The question, why there is also an effect of attention without reward can be overcome by counting the success in fulfilling the task as a reward itself. Thus, either the chance of getting monetary reward or of succeeding in a task could make it more likely that attention will be allocated to the respective stimulus. It is difficult and complicated to more clearly disentangle the influences of attention and reward since they are tightly linked. Furthermore, the use of an objective method limits the set of possible tasks since drifting gratings as a central binocular rivalry task are needed to reliably assess subjective perception. However, future research should further investigate the influences of value and attention on perception in binocular rivalry, for instance, by varying the degree of attention that is needed to fulfill a task and at the same time measuring the effect of reward. Finding a suitable attentional task is the most critical step here. Then the dynamics of the interaction between attention and valuation could be studied in even more detail.

Attentional control

The question as to what the driving factors are that guide the allocation of attention in a natural environment as well as in laboratory settings has been studied for decades by now. Attention can be controlled voluntarily (i.e., goal-driven), or

automatically (i.e., stimulus-driven). This has been formalized in the concept of top-down or bottom-up attentional control, respectively (Posner, 1980; Desimone & Duncan, 1995; Itti & Koch, 2000; Kastner & Ungerleider, 2000). The notion of either task-driven or stimulus-driven processes guiding attentional allocation has recently been extended to also include selection history (Awh et al., 2012). This stands for the prioritization of items that have been attended previously in a given context, which may contrast selection goals. This accounts for results where, for instance, previously selected reward-related features had an effect on perception even though this effect was suboptimal for task performance (Hickey et al., 2010).

In studies 1 and 3 we studied the influences of top-down and bottom-up factors as well as the effect of reward on perception. Since in rivalry the stimulus and thereby the bottom-up factors remains the same throughout the experiment, it is an ideal testbed to investigate the effect of task-dependent influences on perception. In study 1, attention was clearly driven by task-dependent instructions in the attention condition since participants were explicitly told to pay attention to one of the stimuli. This apparently biased competition such that the respective stimulus was seen much longer than the unattended one, but because of the perceptual alternations, competition could not be resolved entirely as discussed earlier. In the reward condition, even though it is not clear to which degree the effect of reward is mediated by attention or vice versa, participants were explicitly instructed which of the stimuli was rewarded so perception was also guided by task-dependent factors. Hence, in study 1 we could confirm that when bottom-up factors stay constant, top-down factors have a strong influence on perception.

To complement the findings on the task-dependent attentional control we induced in study 1, we investigated how a low-level stimulus feature like color influences perception in study 3 and what role attention plays in this context. It is known that color influences human overt visual attention in natural scenes differently depending on color feature (e.g., red-green contrast or saturation) and depicted scene, with particular influence on rainforest and landscape as compared to other scenes (Frey et al., 2008). Although there are even more studies indicating an association between attentional processes and the function of color vision (Maunsell & Treue, 2006; Motter, 1994), color seems to play a minor role in rapid natural scene processing (see section 1.5). Conducting experiments using different SOAs and picture modifications, we shed light into how color influences rapid visual processing. First, we found a small but significant benefit of color over grayscale in our series of RSVP experiments for intermediate (90-100ms) and long (120ms) SOAs. For intermediate presentation durations, color needed to be correct to be beneficial for processing while for longer SOAs original color and modified color pictures were detected equally well. So the effect of color seems to highly depend on presentation duration, and it is likely that the dominant effect accounting for the benefit of color changes with presentation duration from being not beneficial (<=60ms) through being diagnostic for animal pictures (90-100ms) to potentially being beneficial in figure-ground segmentation (120ms). Two of our

findings imply that color influences perception preattentively rather than in an attentional way: color did not aid in categorizing the detected animals, and color did not influence the characteristics of the attentional blink. Thus we could break down the contribution of color to rapid visual processing and its influence on attention in two factors: Diagnosticity and segmentation, dependent on how long it is presented. Importantly, the aformentioned critical SOAs must vary between experiments depending on whether the pictures are presented in an RSVP stream, only followed by a mask or are presented without postmask.

Studies 1 and 3 showed clear dependencies of attentional allocation on either only top-down or only bottom-up factors. But relying on task demands only and ignoring visual conditions or vice versa would be an extremely suboptimal strategy for behavior in real life, so our perception in everyday life must rely on both strategies. Even though in a situation where a task is given almost all eye movements are task-related (Land et al., 1999), in a semi-constrained experiment with a less attention-demanding task more eye movements are explorative ('t Hart & Einhäuser, 2012). Thus in real life, perception and behavior is always guided by both, stimulus salience and task demands, depending on the situation and observer characteristics, like age (Acik et al., 2010) or disease (Dowiasch et al., 2015). In experiment 4 and 5, the task was primarily to first walk along a corridor and after that to track a stationary target on the floor. While in the tracking condition, most eye movements were focused on the target and thus were task-dependent, in the semi-constrained walking condition in the beginning there were many explorative eye movements. Breaking down the single eye movements into task-related and stimulus-related overt visual attention is impossible here. For instance, most downwards pointing eye movements could be considered task-relevant for walking, but participants could also have moved their eyes down to investigate the ground out of interest. Furthermore, during semi-constrained experiments in the real world, internal factors unknown to the experimenter can drive participants' behavior such as interest in the pictures on the wall or the plants. This makes it even harder to clearly disentangle top-down and bottom-up influences on overt visual attention and the discrimination of these factors was outside the scope of study 4 and 5.

There are numerous models capturing aspects of bottom-up attentional control. One of the most influential models describing visual attention in a bottom-up framework is the saliency map (Itti & Koch, 2000; Koch & Ullman, 1985). It combines different visual features (color, orientation, etc.) contributing to attentive selection of a stimulus into a topographical map. At the output stage of the salience map, a WTA circuit is implemented to select the winning location. When it is properly adjusted, it predicts fixation locations well above chance in free viewing paradigms (Harel, Koch, & Perona, 2006). However, it has also been shown that fixations do not necessarily rely on early visual features , but that objects predict fixations better than early salient features (Einhäuser, Spain, & Perona, 2008; Stoll, Thrun, Nuthmann, & Einhäuser, 2015) and task influences

can easily rule out effects of bottom-up features (Einhäuser, Rutishauser, & Koch, 2008). Hence, also top-down influences have to be taken into account to model visual overt attention.

Our model presented in study 2, in contrast, does not rely on specific stimulus characteristics like color or motion but on the in general noisy input originating from the competing stimuli. Thus it is less specific to the exact stimulus used but can more generally cover most varieties of binocular rivalry (face-house, stationary or drifting gratings etc.). Our model receives constant bottom-up input while there is no explicit top-down influence. Hence, the alternations between dominant units are not predicted by the bottom-up information, nor are they explicitly guided by higher-level information. Instead, they are an emerging property of the recurrent connectivity of the network. As described earlier in this section, the properties of our model can also be extended to take into account explicit top-down influences like attention or reward. Hence, our model has proven successful in modeling influences of stimulus strength on perception in rivalry but will also be a good candidate for capturing top-down influences.

Advantages and limitations of real-life eye-tracking

In real-life situations, perception and action are tightly linked. Although not studied explicitly, this coupling manifests itself in gaze allocation during real-life behavior, which naturally includes actions like walking, turning, and target tracking. It was shown that in everyday situations, gaze is highly task-dependent (Land et al., 1999; Land & Hayhoe, 2001), and, when there is no specific task, gaze patterns depend on the environment during natural exploration (Einhäuser, Schumann, et al., 2007). In studies 4 and 5, the gaze pattern and eye-movement behavior of participants was investigated in an everyday situation. Some findings did not correspond with earlier laboratory experiments studying similar participant groups. For instance, in study 4, there was no difference in smooth-pursuit gain between older and younger adults, contradicting earlier laboratory findings (Moschner & Baloh, 1994), and in study 5, the impairment in saccade amplitude and velocity were much less prominent in the real-world measurement than in the fixation task. These findings could be manifestations of the earlier conclusions by 't Hart et al. (2009), that real-world measurements where the whole body moved and laboratory findings, even using highly similar stimuli, lead to different gaze patterns. We do not have the direct comparison of how gaze behavior would have been for the same participants in the laboratory with similar stimuli in our studies. However, in study 5, we could compare data from a standard fixation protocol, which is comparable to standard saccade tasks in the laboratory, to the real-life behavior and saw clear differences in saccadic parameters like peak velocity and amplitude. This could be due to the high number of short saccades in real life as compared to the large saccades in the fixation protocol. Generally, in real life,

saccades are of lower amplitude since head movements can support large gaze shifts, which is impossible in the head-fixed viewing conditions often used in the laboratory. In our study, the two stimuli, the red dots in the fixation protocol and the hallway during walking, were extremely dissimilar. For this reason we cannot conclude how the gaze patterns differ between seeing a similar stimulus while walking or while sitting in a laboratory, as 't Hart et al. (2009) did. However, that remains to be investigated with more age groups and different diseases to complement our findings and to get a deeper insight into what contributions of the eye-movement patterns are due to the stimulus and which characteristics are due to the body movements. But in our studies the goal was to see differences between different age and patient groups in real-life eye movements, not to compare these findings directly to laboratory experiments. Additionally, we wanted to find a measure for differential diagnosis of progressive supranuclear palsy (PSP) and idiopathic Parkinson's diseas (IPD) patients and to investigate how their deficits manifest in real life. Even though there is more research needed on how other diseases that affect eye movements restrict patients in activities of daily living, we succeeded in establishing a method rendering this possible and we can provide initial findings in the case of PSP patients.

The interaction of perception and action is also of high importance with regard to social interactions. Understanding someone else's actions and being able to imitate or react to them is crucial for learning behavior but also for our social status. In an extreme view it might be social interaction that shaped perception, action and cognition (Knoblich & Sebanz, 2006). This makes it even more attractive to study social interaction using action-perception paradigms in real-world settings to investigate how perception and action are influenced by social behavior. First steps have been made into this direction using a mobile eyetracker in human-robot interaction of patients with Autism Spectrum Disorder (Damm et al., 2013). Thus, mobile eye tracking creates new opportunities in investigating action-perception interaction, also in social situations.

Real-life eye-tracking has become an important tool for studying the allocation of visual overt attention in everyday situations (see introduction). This made it possible to test how laboratory findings transfer to real life, which has been studied by 't Hart et al. (2009) and in study 4 and 5 of this thesis. 't Hart et al. recorded eye movements of human observers that viewed the similar visual scenes - one time the scene was looked at in real life, one time a movie of the same scene recorded by a head-fixed camera was presented on a screen. The finding that there are significant differences in viewing patterns between the conditions emphasizes the need for more studies investigating the direct relation between laboratory and real-world findings. Clearly, there are limitations and challenges when investigating eye movements in the real world. Control over the stimulus, which is then the whole field of view, is limited and viewing behavior is more influenced by movements of the body and the head than in most laboratory studies. Furthermore, different classes of eye movements like voluntary saccades and

the reflexive fast phases of the optokinetic nystagmus are difficult and sometimes impossible to distinguish in real-world settings. Thus, in the two last experiments of this thesis, we termed all fast eye movements saccades, irrespective of their reflexive or voluntary nature. And even though the measures we analyzed, like saccade velocity and amplitude, were similar to those that have been studied in laboratory settings, the real-world task the human observers were obliged to do was different from standard procedures used in stationary settings. To make findings better comparable, it could be a next step to conduct standard laboratory experiments with the same subjects that are investigated in the real world and compare eye-movement parameters between conditions and participant groups to test how the procedure affects the parameters of interest and, most importantly, the differences between participant groups. But since the purpose of vision is not to give us an exact representation of the world but to make us able to behave adequately (Einhäuser & König, 2010), studying eye movements while behaving naturally means studying the very nature and purpose of eye movements. Thus, the last two studies of this thesis were one first step in bridging the gap between laboratory findings on the influences of age on vision as well as the impact of PSP on vision and the visual perception during everyday life.

Even though it is definitely interesting to get an insight into how everyday visual perception is affected by age or diseases, the fifth study raised the question, when it is clinically recommendable to use real-life behavior for diagnosis. Since differences between patient groups were most prominent and significant when a standardized fixation protocol was conducted, for a quick diagnosis tool only this part of the experiment is needed. So the advantage of moving around with the eye tracker is not necessarily needed here. However, mobile eye trackers are the most practical device to use in this case since patients can be very immobile and a mobile eye tracker makes it most convenient to measure patients at the bedside. This renders mobile eye tracking an uncomplicated and quick tool for diagnosis.

Concluding, this thesis contributed to the understanding of how specific top-down effects like valuation and attention as well as bottom-up stimulus features like color influence visual perception. I investigated these features in perceptually challenging situations to make transfer to our perception of the real world, which is complex and sometimes ambiguous, easier and more straight-forward. Finally, I measured eye movements in the real world to investigate the impact of age and neurological diseases on eye movements and visual perception. There is still a lot to do in the field of complex-scene perception, or more importantly, real-life perception which is inseparable from behavior itself. Models need to be refined to also capture situations where bottom-up as well as top-down influences affect perception, and more studies are needed to pinpoint and disentangle the mechanisms driving attention in complex situations. Particularly, rather than studying perception and action separately, the interplay between perception and action needs to be studied in more detail to understand this vital relation. My findings

pave the way for further research by showing effects of value on perception and demonstrating the suitability of a WTA model to capture perceptual alternations. Furthermore, by confirming and extending the usability of mobile eye tracking in different participant groups I firmly believe that my thesis may motivate more research, particularly in clinical settings.

References

Acik, A., Sarwary, A., Schultze-Kraft, R., Onat, S., & König, P. (2010). Developmental changes in natural viewing behavior: bottom-up and top-down differences between children, young adults and older adults. *Frontiers in Psychology*, *1*, 207.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464.

Arnold, D. H. (2011). Why is binocular rivalry uncommon? discrepant monocular images in the real world. *Frontiers in Human Neuroscience*, *5*, 116.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*(4), 600–609.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In R. W.A. (Ed.), *Sensory communication* (pp. 217–234). MIT Press Cambridge, MA.

Beets, I. A. M., Rösler, F., Henriques, D. Y. P., Einhäuser, W., & Fiehler, K. (2010). Online action-to-perception transfer: Only percept-dependent action affects perception. *Vision Research*, *50*(24), 2633–2641.

Betz, T., Kietzmann, T. C., Wilming, N., & König, P. (2010). Investigating task-dependent top-down effects on overt visual attention. *Journal of Vision*, *10*(3), 15.

Billino, J., Bremmer, F., & Gegenfurtner, K. R. (2008). Differential aging of motion processing mechanisms: Evidence against general perceptual decline. *Vision Research*, *48*(10), 1254–1261.

Binzegger, T., Douglas, R. J., & Martin, K. A. (2004). A quantitative map of the circuit of cat primary visual cortex. *The Journal of Neuroscience*, *24*(39), 8441–8453.

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind*, *2*(1), 5–38.

Blake, R., Fox, R., & McIntyre, C. (1971). Stochastic properties of stabilized-image binocular rivalry alternations. *Journal of Experimental Psychology*, *88*(3), 327–332.

Blake, R., O'Shea, R. P., & Mueller, T. J. (1992). Spatial zones of binocular rivalry in central and peripheral vision. *Visual Neuroscience*, *8*, 469–478.

Boring, E. G. (1930). A new ambiguous figure. *The American Journal of Psychology*, *42*, 444–445.

Bossink, C., Stalmeier, P., & de Weert, C. M. (1993). A test of levelt's second proposition for binocular rivalry. *Vision Research*, *33*(10), 1413–1419.

Brascamp, J. W., Pearson, J., Blake, R., & van den Berg, A. V. (2009). Intermittent ambiguous stimuli: Implicit memory causes periodic perceptual alternations. *Journal of Vision*, *9*(3), 3.

Brascamp, J. W., van Ee, R., Noest, A. J., Jacobs, R. H. A. H., & van den Berg, A. V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *Journal of Vision*, *6*(11), 1244–1256.

Brascamp, J. W., van Ee, R., Pestman, W. R., & van den Berg, A. V. (2005). Distributions of alternation rates in various forms of bistable perception. *Journal of Vision*, *5*(4), 287–298.

Breese, B. B. (1899). On inhibition. *The Psychological Review: Monograph Supplements*, *3*(1), i.

Bruner, J. S., & Potter, M. C. (1964). Interference in visual recognition. *Science*, *144*(3617), 424–425.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547.

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*(2), 291–328.

Buswell, G. T. (1935). *How people look at pictures*. University of Chicago Press Chicago.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313.

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, *80*(6), 2918–2940.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109–127.

Chun, M. M., & Wolfe, J. M. (2005). Visual attention. In B. E. Goldstein (Ed.), *Blackwell handbook of sensation and perception* (pp. 272–310). Oxford, UK: Blackwell Publishing Ltd.

Damm, O., Malchus, K., Jaecks, P., Krach, S., Paulus, F., Naber, M., ... others (2013). Different gaze behavior in human-robot interaction in asperger's syndrome: An eye-tracking study. In *Ro-man, 2013 ieee* (pp. 368–369).

Deco, G., & Rolls, E. T. (2005). Attention, short-term memory, and action selection: A unifying theory. *Progress in Neurobiology*, *76*(4), 236–256.

Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science*, *20*(6), 778–784.

Della Libera, C., Perlato, A., & Chelazzi, L. (2011). Dissociable effects of reward on attentional learning: From passive associations to active monitoring. *PLoS One*, *6*(4), e19460.

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2010). Key visual features for rapid categorization of animals in natural scenes. *Frontiers in Psychology*, *1*, 21.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837.

Dieter, K. C., & Tadin, D. (2011). Understanding attentional modulation of binocular rivalry: A framework based on biased competition. *Frontiers in Human Neuroscience*, *5*, 155.

Di Pace, E., & Saracini, C. (2014). Action imitation changes perceptual alternations in binocular rivalry. *PloS One*, *9*(5), e98305.

Douglas, R. J., Koch, C., Mahowald, M., Martin, K., & Suarez, H. H. (1995). Recurrent excitation in neocortical circuits. *Science*, *269*(5226), 981–985.

Douglas, R. J., & Martin, K. A. (2004). Neuronal circuits of the neocortex. *Annual Review of Neuroscience*, *27*, 419–451.

Douglas, R. J., & Martin, K. A. (2007). Recurrent neuronal circuits in the neocortex. *Current Biology*, *17*(13), R496–R500.

Dowiasch, S., Backasch, B., Einhäuser, W., Leube, D., Kircher, T., & Bremmer, F. (2015). Eye movements of patients with schizophrenia in a natural environment. *European Archives of Psychiatry and Clinical Neuroscience*, *in press*.

Einhäuser, W., Koch, C., & Makeig, S. (2007). The duration of the attentional blink in natural scenes depends on stimulus category. *Vision Research*, *47*(5), 597–607.

Einhäuser, W., & König, P. (2010). Getting real—sensory processing of natural stimuli. *Current Opinion in Neurobiology*, *20*(3), 389–395.

Einhäuser, W., Martin, K. A., & König, P. (2004). Are switches in perception of the necker cube related to eye position? *European Journal of Neuroscience*, *20*(10), 2811–2818.

Einhäuser, W., Rutishauser, U., & Koch, C. (2008, February). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision*, *8*(2), 2.

Einhäuser, W., Schumann, F., Bardins, S., Bartl, K., Böning, G., Schneider, E., & König, P. (2007). Human eye-head co-ordination in natural exploration. *Network*, *18*(3), 267–297.

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, *8*(14), 18.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, *105*(5), 1704–1709.

Elder, J. H., & Velisavljević, L. (2009). Cue dynamics underlying rapid detection of animals in natural scenes. *Journal of Vision*, *9*(7), 7.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716.

Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? the pragmatic turn in cognitive science. *Trends in Cognitive Sciences*, *17*(5), 202–209.

Eriksen, C. W., & James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*(4), 225–240.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1476–1492.

Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology*, *2*, 243.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, *13*(2), 171–180.

Fairchild, M. D., Johnson, G. M., Babcock, J., & Pelz, J. B. (2001). Is your eye on the ball? eye tracking golfers while putting. *(http://www.cis.rit.edu/people/faculty/fairchild/)*.

Fox, R., & Herrmann, J. (1967). Stochastic properties of binocular rivalry alternations. *Attention, Perception, & Psychophysics*, *2*(9), 432–436.

Freeman, A. W. (2005). Multistage model for binocular rivalry. *Journal of Neurophysiology*, *94*(6), 4412–4420.

Frey, H.-P., Honey, C., & König, P. (2008). What's color got to do with it? the influence of color on visual attention in different categories. *Journal of Vision*, *8*(14), 6.

Gallant, J. L., Connor, C. E., & van Essen, D. C. (1998). Neural activity in areas v1, v2 and v4 during free viewing of natural scenes compared to controlled viewing. *Neuroreport*, *9*(9), 2153–2158.

Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, *10*(13), 805–808.

Haegerstrom-Portnoy, G. (2005). The glenn a. fry award lecture 2003: Vision in elders—summary of findings of the ski study. *Optometry & Vision Science*, *82*(2), 87–93.

Hahnloser, R., Douglas, R. J., Mahowald, M., & Hepp, K. (1999). Feedback interactions between neuronal pointers and maps for attentional processing. *Nature Neuroscience*, *2*(8), 746–752.

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. In *Advances in neural information processing systems* (pp. 545–552).

Hayhoe, M., Mennie, N., Sullivan, B., & Gorgos, K. (2002). The role of internal models and prediction in catching balls. In *Proceedings of the american association for artificial intelligence.*

Hecht, H., Vogt, S., & Prinz, W. (2001). Motor learning enhances perceptual judgment: A case for action-perception transfer. *Psychological Research*, *65*(1), 3–14.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *The Journal of Neuroscience*, *30*(33), 11096–11103.

Irving, E. L., Steinbach, M. J., Lillakas, L., Babu, R. J., & Hutchings, N. (2006). Horizontal saccade dynamics across the human life span. *Investigative Ophthalmology & Visual Science*, *47*(6), 2478–2484.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10), 1489–1506.

James, W. (1890). *The principles of psychology.* New York: Holt.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*, *23*(1), 315–341.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776.

Klein, R., Klein, B., Linton, K., & De Mets, D. L. (1991). The beaver dam eye study: Visual acuity. *Ophthalmology*, *98*(8), 1310–1315.

Klink, P. C., van Ee, R., & van Wezel, R. J. A. (2008). General validity of levelt's propositions reveals common computational mechanisms for visual rivalry. *PLoS One*, *3*(10), e3473.

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, *15*(3), 99–104.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–227.

Lago-Fernández, L. F., & Deco, G. (2002). A model of binocular rivalry based on competition in it. *Neurocomputing*, *44*, 503–507.

Lago-Rodríguez, A., Cheeran, B., Koch, G., Hortobagy, T., & Fernandez-del-Olmo, M. (2014). The role of mirror neurons in observational motor learning: an integrative review. *European Journal of Human Movement*, *32*, 82–103.

Laing, C. R., & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience*, *12*(1), 39–53.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25-26), 3559–3565.

Land, M. F., & Lee, D. (1994). Where we look when we steer. *Nature*, *369*(6483), 742–744.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience*, *3*(12), 1340–1345.

Land, M. F., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, *28*(11), 1311–1328.

Land, M. F., & Tatler, B. W. (2001). Steering with the head: The visual strategy of a racing driver. *Current Biology*, *11*(15), 1215–1220.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience*, *2*(4), 375–381.

Leeper, R. (1935). A study of a neglected portion of the field of learning—the development of sensory organization. *Journal of Genetic Psychology*, *46*(1), 41–75.

Leigh, R. J., & Kennard, C. (2004). Using saccades as a research tool in the clinical neurosciences. *Brain*, *127*(3), 460–477.

Leigh, R. J., & Zee, D. S. (2006). *The neurology of eye movements*. Oxford University Press.

Lennie, P. (2003). The cost of cortical computation. *Current Biology*, *13*(6), 493–497.

Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*(6565), 549–553.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254–264.

Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, *5*(6), 605–609.

Levelt, W. J. M. (1967). Note on the distribution of dominance times in binocular rivalry. *British Journal of Psychology*, *58*(1-2), 143–145.

Levelt, W. J. M. (1968). *On binocular rivalry*. The Hague: Mouton.

Li, F. F., van Rullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596–9601.

Lich, M., & Bremmer, F. (2014). Self-motion perception in the elderly. *Frontiers in Human Neuroscience*, *8*, 681.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. MIT Press Cambridge, MA.

MacLean, M. H., & Arnell, K. M. (2012). A conceptual and methodological framework for measuring and modulating the attentional blink. *Attention, Perception, & Psychophysics*, *74*(6), 1080–1097.

Madden, D. J. (2007). Aging and visual attention. *Current Directions in Psychological Science*, *16*(2), 70–74.

Maunsell, J. H. R., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual

area (mt) of the macaque monkey. *The Journal of Neuroscience*, *10*(10), 3323–3334.

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322.

Mazer, J. A., & Gallant, J. L. (2003). Goal-related activity in v4 during free viewing visual search: Evidence for a ventral stream visual salience map. *Neuron*, *40*(6), 1241–1250.

McMains, S., & Kastner, S. (2009). Visual attention. In M. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedia of neuroscience* (p. 4296-4302). Springer Berlin Heidelberg.

Meng, M., & Potter, M. C. (2008). Detecting and remembering pictures with and without visual noise. *Journal of vision*, *8*(9), 7.

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? differences between binocular rivalry and ambiguous figures. *Journal of Vision*, *4*(7), 2.

Moreno-Bote, R., Rinzel, J., & Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *Journal of Neurophysiology*, *98*(3), 1125–1139.

Moreno-Bote, R., Shpiro, A., Rinzel, J., & Rubin, N. (2010). Alternation rate in perceptual bistability is maximal at and symmetric around equi-dominance. *Journal of Vision*, *10*(11), 1.

Moschner, C., & Baloh, R. W. (1994). Age-related changes in visual tracking. *Journal of Gerontology*, *49*(5), M235–M238.

Motter, B. C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area v4. *The Journal of Neuroscience*, *14*(4), 2178–2189.

Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research*, *121*(4), 391–400.

Naber, M., Frässle, S., & Einhäuser, W. (2011). Perceptual rivalry: reflexes reveal the gradual nature of visual awareness. *PloS One*, *6*(6), e20910.

Naber, M., Gruenhage, G., & Einhäuser, W. (2010). Tri-stable stimuli reveal interactions among subsequent percepts: Rivalry is biased by perceptual history. *Vision Research*, *50*(8), 818–828.

Necker, L. A. (1832). Lxi. observations on some remarkable optical phaeomena seen in switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid. *The London and Edinburgh Philosophy Magazine and Journal of Science*, *3*, 329–337.

Noest, A. J., van Ee, R., Nijs, M. M., & van Wezel, R. J. A. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: A low-level neural model. *Journal of Vision*, *7*(8).

Nusseck, M., Lagarde, J., Bardy, B., Fleming, R., & Bülthoff, H. H. (2007). Perception and prediction of simple object interactions. In *Proceedings of*

the 4th symposium on applied perception in graphics and visualization (pp. 27–34).

Ooi, T. L., & He, Z. J. (1999). Binocular rivalry and visual awareness: The role of attention. *Perception*, *28*(5), 551–574.

Orbach, J., Ehrlich, D., & Helen, A. H. (1963). Reversibility of the necker cube: I. an examination of the concept "satiation of orientation". *Perceptual and Motor Skills*, *17*(2), 439–458.

Paffen, C. L. E., & Alais, D. (2011). Attentional modulation of binocular rivalry. *Frontiers in Human Neuroscience*, *5*, 105.

Paffen, C. L. E., Alais, D., & Verstraten, F. A. (2006). Attention speeds binocular rivalry. *Psychological Science*, *17*(9), 752–756.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology* (Vol. 1). MIT press Cambridge, MA.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, *81*(1), 10–15.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an rsvp task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1), 31–40.

Rothkirch, M., Schmack, K., Deserno, L., Darmohray, D., & Sterzer, P. (2014). Attentional modulation of reward processing in the human brain. *Human brain mapping*, *35*(7), 3036–3051.

Rousselet, G. A., Macé, M. J.-M., & Fabre-Thorpe, M. (2003). Is it an animal? is it a human face? fast processing in upright and inverted natural scenes. *Journal of Vision*, *3*(6), 5.

Rubin, E. (1915). *Synsoplevede figurer: studier i psykologisk analyse. 1. del.* Gyldendalske Boghandel, Nordisk Forlag.

Rutishauser, U., & Douglas, R. J. (2009). State-dependent computation using coupled recurrent networks. *Neural Computation*, *21*(2), 478–509.

Rutishauser, U., Douglas, R. J., & Slotine, J.-J. (2011). Collective stability of networks of winner-take-all circuits. *Neural Computation*, *23*(3), 735–773.

Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences*, *11*(8), 349–355.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological Science*, *5*(4), 195–200.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61*(5), 700–707.

Sekuler, R., & Blake, R. (1985). *Perception*. New York: Alfred A. Knopf.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074.

Stănişor, L., van der Togt, C., Pennartz, C. M., & Roelfsema, P. R. (2013). A unified selection signal for attention and reward in primary visual cortex. *Proceedings of the National Academy of Sciences*, *110*(22), 9136–9141.

Stoll, J., Thrun, M., Nuthmann, A., & Einhäuser, W. (2015). Overt attention in natural scenes: Objects dominate features. *Vision Research*, *107*, 36–48.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network*, *14*(3), 391–412.

Treue, S., & Martinez-Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*(6736), 575–579.

Treue, S., & Maunsell, J. H. R. (1996). Attentional modulation of visual motion processing in cortical areas mt and mst. *Nature*, *382*(6591), 539–541.

van Dam, L. C. J., & van Ee, R. (2006). The role of saccades in exerting voluntary control in perceptual and binocular rivalry. *Vision research*, *46*(6), 787–799.

van Ee, R. (2009). Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: Evidence from serial correlations in perceptual bistability. *Journal of the Optical Society of America*, *26*(12), 2612–2622.

van Ee, R., van Dam, L. C. J., & Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Research*, *45*(1), 41–55.

van Rullen, R., & Thorpe, S. J. (2001). Is it a bird? is it a plane? ultra-rapid visual categorisation of natural and artifactual objects. *Perception*, *30*(6), 655–668.

von Helmholtz, H. (1867). *Handbuch der physiologischen optik: mit 213 in den text eingedruckten holzschnitten und 11 tafeln*. Voss, Leipzig.

Wheatstone, C. (1838). Contributions to the physiology of vision.–part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical transactions of the Royal Society of London*, *128*, 371–394.

Whiting, W. L., Madden, D. J., Pierce, T. W., & Allen, P. A. (2005). Searching from the top down: Ageing and attentional guidance during singleton detection. *The Quarterly Journal of Experimental Psychology Section A*, *58*(1), 72–97.

Wichmann, F. A., Braun, D. I., & Gegenfurtner, K. R. (2006). Phase noise and the classification of natural images. *Vision Research*, *46*(8), 1520–1529.

Wichmann, F. A., Drewes, J., Rosas, P., & Gegenfurtner, K. R. (2010). Animal detection in natural scenes: Critical features revisited. *Journal of Vision*, *10*(4), 6.

Wilson, H. R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proceedings of the National Academy of Sciences*, *100*(24), 14499–14503.

Wilson, H. R. (2005). Rivalry and perceptual oscillations: A dynamical synthesis. In D. Alais & R. Blake (Eds.), *Binocular rivalry* (pp. 317–335). MIT Press Cambridge, MA.

Yao, A. Y., & Einhäuser, W. (2008). Color aids late but not early stages of rapid natural scene recognition. *Journal of Vision*, *8*(16), 12.

Yarbus, A. L. (1967). *Eye movements and vision.* New York: Plenum Press.

Yuille, A., & Geiger, D. (2003). Winner-take-all networks. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 1228–1231). MIT Press Cambridge, MA.

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *in press*.

't Hart, B. M., & Einhäuser, W. (2012). Mind the step: Complementary effects of an implicit task on eye and head movements in real-life gaze allocation. *Experimental Brain Research*, *223*(2), 233–249.

't Hart, B. M., Vockeroth, J., Schumann, F., Bartl, K., Schneider, E., König, P., & Einhäuser, W. (2009). Gaze allocation in natural stimuli: Comparing free exploration to head-fixed viewing conditions. *Visual Cognition*, *17*(6/7), 1132–1158.

Declaration of the authors' contributions to the studies

The main part of this thesis consists of five articles I co-authored, which are published or accepted for publication.

Specifically, chapter 2.1 is published as

> **Marx S**, *& Einhäuser W (2015). Reward modulates perception in binocular rivalry. Journal of Vision, 15(1):11, 1-13.*

Authors' contributions: **SM** and WE conceived the study; **SM** designed the experiment, performed data collection and data analysis; **SM** and WE wrote the article.

Chapter 2.2 is published as

> **Marx S**,*Gruenhage G, Walper D, Rutishauser U, & Einhäuser W (2015). Competition with and without priority control: linking rivalry to attention through winner-take-all networks with memory. Annals of the New York Academy of Sciences, 1339, 138-153.*

Authors' contributions: **SM**, GG, UR and WE conceived the study; **SM**, GG, and UR conducted the modeling; **SM** and DW designed and conducted the experiment, and analyzed the experimental data; **SM**, UR and WE wrote the paper.

Chapter 2.3 is published as

> **Marx S**, *Hansen-Goos O, Thrun M, & Einhäuser W (2014). Rapid serial processing of natural scenes: Color modulates detection but neither recognition nor the attentional blink. Journal of Vision, 14(14):4, 1-18.*

Authors' contributions: **SM** and WE conceived the study; **SM**, OH, MT, WE designed the experiment and conducted data collection; **SM** and WE analyzed the data and wrote the article.

Chapter 2.4 is published as

> *Dowiasch S.,* **Marx S**, *Einhäuser W, & Bremmer F (2015). Effects of aging on eye movements in the real world. Frontiers in Human Neuroscience, 9(46).*

Authors' contributions: SD and FB conceived the study; SD and **SM** collected and analyzed the data; SD wrote the article; WE and FB supervised data analysis and all authors proof-read the manuscript.

Chapter 2.5 is published as

> **Marx S**, *Respondek G*, Stamelou M, Dowiasch S, Stoll J, Bremmer F, Oertel W H, Höglinger G U**, & Einhäuser W** (2012). Validation of mobile eye-tracking as novel and efficient means for differentiating progressive supranuclear palsy from Parkinson's disease. Frontiers in Behavioral Neuroscience, 6(88).*

Authors' contributions: **SM**, MS, FB, WHO, GUH and WE conceived the study; GR, MS, WHO and GUH gave access to patients; **SM**, GR, MS and SD collected the data; JS provided technical support; **SM**, SD and JS analyzed the data; GR, **SM** and WE wrote the article, all authors proof-read the article. Parts of the data were included in SM's diploma thesis. During the PhD, substantially more data were collected, all data were reanalyzed and the manuscript was written.

* ** denote equal contribution

(Svenja Marx) (Prof. Wolfgang Einhäuser-Treyer)

Curriculum Vitae

12/2011 - 04/2015	PhD-Student at the Philipps-University Marburg, AG Neurophysik (Prof. Dr. Wolfgang Einhäuser-Treyer)
10/2011	Diploma in Physics
10/2006 - 10/2011	Study of Physics at the Philipps-University Marburg
06/2005	Abitur at the Weidigschule Butzbach

Danksagungen

Hiermit möchte ich mich von ganzem Herzen bei allen Menschen bedanken, die direkt oder indirekt an der Entstehung und dem erfolgreichen Abschluss dieser Doktorarbeit beteiligt waren.

An erster Stelle danke ich Prof. Wolfgang Einhäuser-Treyer, der mich angeleitet, motiviert, inspiriert und ständig herausgefordert hat. Danke für die Freiheit, mit der ich eigene Ideen verfolgen konnte und für Ermutigung in frustrierenden Zeiten.

Bei Ueli Rutishauser bedanke ich mich für die gute Zusammenarbeit in Studie 2 und dafür, dass er mir eine spannende und lehrreiche Zeit in LA ermöglicht hat.

Dank gilt auch meinem Zweitgutachter, Prof. Peter König, und den anderen Mitgliedern meiner Prüfungskommission, Prof. Heinz Jänsch und Prof. Reinhard Noack, für ihre Zeit und ihr Interesse.

Ohne meine Kollegen und Kolleginnen hätte das Forschen nicht halb so viel Freude gemacht und wären einige Projekte nicht möglich gewesen. Dank gebührt hier vor allem Stefan Dowiasch, der mich zu Beginn meiner Zeit in der AG Neurophysik sehr unterstützt hat und über die Jahre ein hilfreicher Diskussionspartner und angenehmer Büropartner war. Danke auch an Josef Stoll, Alexander Platzner und Daniel Walper für die (mess-) technische Unterstützung meiner Arbeit sowie Frau Thomas für ihre unersetzbare Unterstützung in allen organisatorischen Fragen.

Die vielen Versuchspersonen, die an zum Teil extrem anstrengenden Experimenten teilgenommen haben, ermöglichten die Datenaufnahme und damit die Grundlage meiner Arbeit. Vielen Dank für euer Engagement.

Ganz besonders möchte ich mich bei meiner Familie bedanken, die mich unterstützt und immer an mich geglaubt hat. Danke für eure Liebe, euer Verständnis, euer Interesse und die Kraft, die ihr mir in den letzten Jahren gegeben habt.

Dank gebührt auch meinen Freunden in Marburg, Butzbach und dem Rest der Welt für fachliche, besonders aber persönliche Unterstützung in allen Fragen des Doktorandenlebens und darüber hinaus.

Steffen, ich danke dir für deine unermüdliche Unterstützung auf den letzten Metern und davor. Ich freue mich auf alles, was kommt.

Ich danke dem Vater des Lichts, der Grundlage und Durchführung dieser Arbeit erst möglich gemacht hat.